FUNDAMENTAL METHODS OF MATHEMATICAL ECONOMICS

Alpha C. Chiang

FUNDAMENTAL METHODS OF MATHEMATICAL ECONOMICS

Third Edition

Alpha C. Chiang

Professor of Economics
The University of Connecticut

McGraw-Hill, Inc.

New York St. Louis San Francisco Auckland Bogotá
Caracas Lisbon London Madrid Mexico City Milan
Montreal New Delhi San Juan Singapore
Sydney Tokyo Toronto

This book was set in Times Roman by Science Typographers, Inc.
The editors were Patricia A. Mitchell and Gail Gavert;
the production supervisor was Leroy A. Young.
The cover was designed by Carla Bauer.

FUNDAMENTAL METHODS OF MATHEMATICAL ECONOMICS

21 22 23 24 25 26 27 28 29 30 FGRFGR 9 9

ISBN 0-07-010813-7

Library of Congress Cataloging in Publication Data

Chiang, Alpha C., date
Fundamental methods of mathematical economics.

Bibliography: p.
Includes index.
1. Economics, Mathematical. I. Title.
HB135.C47 1984 330'.01'51 83-19609
ISBN 0-07-010813-7

This book is printed on acid-free paper.

CONTENTS

PART ONE

INTRODUCTION

CHAPTER

ONE

THE NATURE OF MATHEMATICAL ECONOMICS

Mathematical economics is not a distinct branch of economics in the sense that public finance or international trade is. Rather, it is an *approach* to economic analysis, in which the economist makes use of mathematical symbols in the statement of the problem and also draws upon known mathematical theorems to aid in reasoning. As far as the specific subject matter of analysis goes, it can be micro- or macroeconomic theory, public finance, urban economics, or what not.

Using the term *mathematical economics* in the broadest possible sense, one may very well say that every elementary textbook of economics today exemplifies mathematical economics insofar as geometrical methods are frequently utilized to derive theoretical results. Conventionally, however, mathematical economics is reserved to describe cases employing mathematical techniques beyond simple geometry, such as matrix algebra, differential and integral calculus, differential equations, difference equations, etc. It is the purpose of this book to introduce the reader to the most fundamental aspects of these mathematical methods—those encountered daily in the current economic literature.

1.1 MATHEMATICAL VERSUS NONMATHEMATICAL ECONOMICS

Since mathematical economics is merely an approach to economic analysis, it should not and does not differ from the *non*mathematical approach to economic analysis in any fundamental way. The purpose of any theoretical analysis, regardless of the approach, is always to derive a set of conclusions or theorems from a given set of assumptions or postulates via a process of reasoning. The major difference between "mathematical economics" and "literary economics" lies principally in the fact that, in the former, the assumptions and conclusions are stated in mathematical symbols rather than words and in equations rather than sentences; moreover, in place of literary logic, use is made of mathematical theorems—of which there exists an abundance to draw upon—in the reasoning process. Inasmuch as symbols and words are really equivalents (witness the fact that symbols are usually defined in words), it matters little which is chosen over the other. But it is perhaps beyond dispute that symbols are more convenient to use in deductive reasoning, and certainly are more conducive to conciseness and preciseness of statement.

The choice between literary logic and mathematical logic, again, is a matter of little import, but mathematics has the advantage of forcing analysts to make their assumptions explicit at every stage of reasoning. This is because mathematical theorems are usually stated in the "if-then" form, so that in order to tap the "then" (result) part of the theorem for their use, they must first make sure that the "if" (condition) part does conform to the explicit assumptions adopted.

Granting these points, though, one may still ask why it is necessary to go beyond geometric methods. The answer is that while geometric analysis has the important advantage of being visual, it also suffers from a serious dimensional limitation. In the usual graphical discussion of indifference curves, for instance, the standard assumption is that only *two* commodities are available to the consumer. Such a simplifying assumption is not willingly adopted but is forced upon us because the task of drawing a three-dimensional graph is exceedingly difficult and the construction of a four- (or higher) dimensional graph is actually a physical impossibility. To deal with the more general case of 3, 4, or n goods, we must instead resort to the more flexible tool of equations. This reason alone should provide sufficient motivation for the study of mathematical methods beyond geometry.

In short, we see that the mathematical approach has claim to the following advantages: (1) The "language" used is more concise and precise; (2) there exists a wealth of mathematical theorems at our service; (3) in forcing us to state explicitly all our assumptions as a prerequisite to the use of the mathematical theorems, it keeps us from the pitfall of an unintentional adoption of unwanted implicit assumptions; and (4) it allows us to treat the general n-variable case.

Against these advantages, one sometimes hears the criticism that a mathematically derived theory is inevitably *unrealistic*. However, this criticism is not valid. In fact, the epithet "unrealistic" cannot even be used in criticizing economic theory in general, whether or not the approach is mathematical. Theory is by its very nature an abstraction from the real world. It is a device for singling out only the most essential factors and relationships so that we can study the crux of the problem at hand, free from the many complications that do exist in the actual world. Thus the statement "theory lacks realism" is merely a truism that cannot be accepted as a valid criticism of theory. It then follows logically that it is quite meaningless to pick out any one approach to theory as "unrealistic." For example, the theory of firm under pure competition is unrealistic, as is the theory of firm under imperfect competition, but whether these theories are derived mathematically or not is irrelevant and immaterial.

In sum, we might liken the mathematical approach to a "mode of transportation" that can take us from a set of postulates (point of departure) to a set of conclusions (destination) at a good speed. Common sense would tell us that, if you intend to go to a place 2 miles away, you will very likely prefer driving to walking, unless you have time to kill or want to exercise your legs. Similarly, as a theorist who wishes to get to your conclusions more rapidly, you will find it convenient to "drive" the vehicle of mathematical techniques appropriate for your particular purpose. You will, of course, have to take "driving lessons" first; but since the skill thus acquired tends to be of service for a long, long while, the time and effort required would normally be well spent indeed.

For a serious "driver"—to continue with the metaphor—some solid lessons in mathematics are imperative. It is obviously impossible to introduce all the mathematical tools used by economists in a single volume. Instead, we shall concentrate on only those that are mathematically the most fundamental and economically the most relevant. Even so, if you work through this book conscientiously, you should at least become proficient enough to comprehend most of the professional articles you will come across in such periodicals as the *American Economic Review, Quarterly Journal of Economics, Journal of Political Economy, Review of Economics and Statistics, and Economic Journal.* Those of you who, through this exposure, develop a serious interest in mathematical economics can then proceed to a more rigorous and advanced study of mathematics.

1.2 MATHEMATICAL ECONOMICS VERSUS ECONOMETRICS

The term "mathematical economics" is sometimes confused with a related term, "econometrics." As the "metric" part of the latter term implies, econometrics is concerned mainly with the measurement of economic data. Hence it deals with the study of *empirical* observations using statistical methods of estimation and hypothesis testing. Mathematical economics, on the other hand, refers to the application of mathematics to the purely *theoretical* aspects of economic analysis, with little or no concern about such statistical problems as the errors of measurement of the variables under study.

In the present volume, we shall confine ourselves to mathematical economics. That is, we shall concentrate on the application of mathematics to deductive reasoning rather than inductive study, and as a result we shall be dealing primarily with theoretical rather than empirical material. This is, of course, solely a matter of choice of the scope of discussion, and it is by no means implied that econometrics is less important.

Indeed, empirical studies and theoretical analyses are often complementary and mutually reinforcing. On the one hand, theories must be tested against empirical data for validity before they can be applied with confidence. On the other, statistical work needs economic theory as a guide, in order to determine the most relevant and fruitful direction of research. A classic illustration of the complementary nature of theoretical and empirical studies is found in the study of the aggregate consumption function. The theoretical work of Keynes on the consumption function led to the statistical estimation of the propensity to consume, but the statistical findings of Kuznets and Goldsmith regarding the relative long-run constancy of the propensity to consume (in contradiction to what might be expected from the Keynesian theory), in turn, stimulated the refinement of aggregate consumption theory by Duesenberry, Friedman, and others.*

In one sense, however, mathematical economics may be considered as the more basic of the two: for, to have a meaningful statistical and econometric study, a good theoretical framework—preferably in a mathematical formulation—is indispensable. Hence the subject matter of the present volume should be useful not only for those interested in theoretical economics, but also for those seeking a foundation for the pursuit of econometric studies.

* John M. Keynes, *The General Theory of Employment, Interest and Money*, Harcourt, Brace and Company, Inc., New York, 1936, Book III; Simon Kuznets, *National Income: A Summary of Findings*, National Bureau of Economic Research, 1946, p. 53; Raymond Goldsmith, *A Study of Saving in the United States*, vol. I, Princeton University Press, Princeton, N.J., 1955, chap. 3; James S. Duesenberry, *Income, Saving, and the Theory of Consumer Behavior*, Harvard University Press, Cambridge, Mass., 1949; Milton Friedman, *A Theory of the Consumption Function*, National Bureau of Economic Research, Princeton University Press, Princeton, N.J., 1957.

CHAPTER TWO
ECONOMIC MODELS

As mentioned before, any economic theory is necessarily an abstraction from the real world. For one thing, the immense complexity of the real economy makes it impossible for us to understand all the interrelationships at once; nor, for that matter, are all these interrelationships of equal importance for the understanding of the particular economic phenomenon under study. The sensible procedure is, therefore, to pick out what appeal to our reason to be the primary factors and relationships relevant to our problem and to focus our attention on these alone. Such a deliberately simplified analytical framework is called an *economic model*, since it is only a skeletal and rough representation of the actual economy.

2.1 INGREDIENTS OF A MATHEMATICAL MODEL

An economic model is merely a theoretical framework, and there is no inherent reason why it must be mathematical. If the model *is* mathematical, however, it will usually consist of a set of *equations* designed to describe the structure of the model. By relating a number of *variables* to one another in certain ways, these equations give mathematical form to the set of analytical assumptions adopted. Then, through application of the relevant mathematical operations to these equations, we may seek to derive a set of conclusions which logically follow from those assumptions.

Variables, Constants, and Parameters

A *variable* is something whose magnitude can change, i.e., something that can take on different values. Variables frequently used in economics include price, profit, revenue, cost, national income, consumption, investment, imports, exports, and so on. Since each variable can assume various values, it must be represented by a symbol instead of a specific number. For example, we may represent price by P, profit by π, revenue by R, cost by C, national income by Y, and so forth. When we write $P = 3$ or $C = 18$, however, we are "freezing" these variables at specific values (in appropriately chosen units).

Properly constructed, an economic model can be solved to give us the *solution values* of a certain set of variables, such as the market-clearing level of price, or the profit-maximizing level of output. Such variables, whose solution values we seek from the model, are known as *endogenous variables* (originating from within). However, the model may also contain variables which are assumed to be determined by forces external to the model, and whose magnitudes are accepted as given data only; such variables are called *exogenous variables* (originating from without). It should be noted that a variable that is endogenous to one model may very well be exogenous to another. In an analysis of the market determination of wheat price (P), for instance, the variable P should definitely be endogenous; but in the framework of a theory of consumer expenditure, P would become instead a datum to the individual consumer, and must therefore be considered exogenous.

Variables frequently appear in combination with fixed numbers or constants, such as in the expressions $7P$ or $0.5R$. A *constant* is a magnitude that does not change and is therefore the antithesis of a variable. When a constant is joined to a variable, it is often referred to as the *coefficient* of that variable. However, a coefficient may be symbolic rather than numerical. We can, for instance, let the symbol a stand for a given constant and use the expression aP in lieu of $7P$ in a model, in order to attain a higher level of generality (see Sec. 2.7). This symbol a is a rather peculiar case—it is supposed to represent a given constant, and yet, since we have not assigned to it a specific number, it can take virtually any value. In short, it is a *constant* that is *variable*! To identify its special status, we give it the distinctive name *parametric constant* (or simply *parameter*).

It must be duly emphasized that, although different values can be assigned to a parameter, it is nevertheless to be regarded as a datum in the model. It is for this reason that people sometimes simply say "constant" even when the constant is parametric. In this respect, parameters closely resemble exogenous variables, for both are to be treated as "givens" in a model. This explains why many writers, for simplicity, refer to both collectively with the single designation "parameters."

As a matter of convention, parametric constants are normally represented by the symbols a, b, c, or their counterparts in the Greek alphabet: α, β, and γ. But other symbols naturally are also permissible. As for exogenous variables, in order that they can be visually distinguished from their endogenous cousins, we shall follow the practice of attaching a subscript 0 to the chosen symbol. For example, if P symbolizes price, then P_0 signifies an exogenously determined price.

Equations and Identities

Variables may exist independently, but they do not really become interesting until they are related to one another by equations or by inequalities. At this juncture we shall discuss equations only.

In economic applications we may distinguish between three types of equation: definitional equations, behavioral equations, and equilibrium conditions.

A *definitional equation* sets up an identity between two alternate expressions that have exactly the same meaning. For such an equation, the identical-equality sign $\equiv$ (read: "is identically equal to") is often employed in place of the regular equals sign $=$, although the latter is also acceptable. As an example, total profit is defined as the excess of total revenue over total cost; we can therefore write

$$\pi \equiv R - C$$

A *behavioral equation*, on the other hand, specifies the manner in which a variable behaves in response to changes in other variables. This may involve either human behavior (such as the aggregate consumption pattern in relation to national income) or nonhuman behavior (such as how total cost of a firm reacts to output changes). Broadly defined, behavioral equations can be used to describe the general institutional setting of a model, including the technological (e.g., production function) and legal (e.g., tax structure) aspects. Before a behavioral equation can be written, however, it is always necessary to adopt definite assumptions regarding the behavior pattern of the variable in question. Consider the two cost functions

$$C = 75 + 10Q \tag{2.1}$$

$$C = 110 + Q^2 \tag{2.2}$$

where Q denotes the quantity of output. Since the two equations have different forms, the production condition assumed in each is obviously different from the other. In (2.1), the fixed cost (the value of C when $Q = 0$) is 75, whereas in (2.2) it is 110. The variation in cost is also different. In (2.1), for each unit increase in Q, there is a constant increase of 10 in C. But in (2.2), as Q increases unit after unit, C will increase by progressively larger amounts. Clearly, it is primarily through the specification of the form of the behavioral equations that we give mathematical expression to the assumptions adopted for a model.

The third type of equations, *equilibrium conditions*, have relevance only if our model involves the notion of equilibrium. If so, the equilibrium condition is an equation that describes the prerequisite for the attainment of equilibrium. Two of the most familiar equilibrium conditions in economics are

$$Q_d = Q_s \qquad \text{[quantity demanded = quantity supplied]}$$

and

$$S = I \qquad \text{[intended saving = intended investment]}$$

which pertain, respectively, to the equilibrium of a market model and the equilibrium of the national-income model in its simplest form. Because equations of this type are neither definitional nor behavioral, they constitute a class by themselves.

2.2 THE REAL-NUMBER SYSTEM

Equations and variables are the essential ingredients of a mathematical model. But since the values that an economic variable takes are usually numerical, a few words should be said about the number system. Here, we shall deal only with so-called "real numbers."

Whole numbers such as 1, 2, 3, ... are called *positive integers*; these are the numbers most frequently used in counting. Their negative counterparts $-1, -2, -3, \ldots$ are called *negative integers*; these can be employed, for example, to indicate subzero temperatures (in degrees). The number 0 (zero), on the other hand, is neither positive nor negative, and is in that sense unique. Let us lump all the positive and negative integers and the number zero into a single category, referring to them collectively as the *set of all integers*.

Integers, of course, do not exhaust all the possible numbers, for we have *fractions*, such as $\frac{2}{3}$, $\frac{5}{4}$, and $\frac{7}{3}$, which—if placed on a ruler—would fall between the integers. Also, we have negative fractions, such as $-\frac{1}{2}$ and $-\frac{2}{5}$. Together, these make up the *set of all fractions*.

The common property of all fractional numbers is that each is expressible as a ratio of two integers; thus fractions qualify for the designation *rational numbers* (in this usage, rational means *ratio*-nal). But integers are also rational, because any integer n can be considered as the ratio $n/1$. The set of all integers and the set of all fractions together form the *set of all rational numbers*.

Once the notion of rational numbers is used, however, there naturally arises the concept of *irrational numbers*—numbers that *cannot* be expressed as ratios of a pair of integers. One example is the number $\sqrt{2} = 1.4142\ldots$, which is a nonrepeating, nonterminating decimal. Another is the special constant $\pi = 3.1415\ldots$ (representing the ratio of the circumference of any circle to its diameter), which is again a nonrepeating, nonterminating decimal, as is characteristic of all irrational numbers.

Each irrational number, if placed on a ruler, would fall between two rational numbers, so that, just as the fractions fill in the gaps between the integers on a ruler, the irrational numbers fill in the gaps between rational numbers. The result of this filling-in process is a continuum of numbers, all of which are so-called "real numbers." This continuum constitutes the *set of all real numbers*, which is often denoted by the symbol R. When the set R is displayed on a straight line (an extended ruler), we refer to the line as the *real line*.

In Fig. 2.1 are listed (in the order discussed) all the number sets, arranged in relationship to one another. If we read from bottom to top, however, we find in effect a classificatory scheme in which the set of real numbers is broken down into its component and subcomponent number sets. This figure therefore is a summary of the structure of the real-number system.

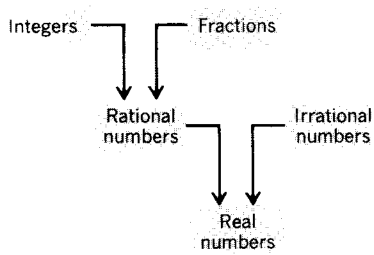

Figure 2.1

Real numbers are all we need for the first 14 chapters of this book, but they are not the only numbers used in mathematics. In fact, the reason for the term "real" is that there are also "imaginary" numbers, which have to do with the square roots of negative numbers. That concept will be discussed later, in Chap. 15.

2.3 THE CONCEPT OF SETS

We have already employed the word "set" several times. Inasmuch as the concept of sets underlies every branch of modern mathematics, it is desirable to familiarize ourselves at least with its more basic aspects.

Set Notation

A *set* is simply a collection of distinct objects. These objects may be a group of (distinct) numbers, or something else. Thus, all the students enrolled in a particular economics course can be considered a set, just as the three integers 2, 3, and 4 can form a set. The objects in a set are called the *elements* of the set.

There are two alternative ways of writing a set: by *enumeration* and by *description*. If we let S represent the set of three numbers 2, 3, and 4, we can write, by enumeration of the elements,

$$S = \{2, 3, 4\}$$

But if we let I denote the set of *all* positive integers, enumeration becomes difficult, and we may instead simply describe the elements and write

$$I = \{x \mid x \text{ a positive integer}\}$$

which is read as follows: "I is the set of all (numbers) x, such that x is a positive integer." Note that braces are used to enclose the set in both cases. In the descriptive approach, a vertical bar (or a colon) is always inserted to separate the general symbol for the elements from the description of the elements. As another example, the set of all real numbers greater than 2 but less than 5 (call it J) can be expressed symbolically as

$$J = \{x \mid 2 < x < 5\}$$

Here, even the descriptive statement is symbolically expressed.

A set with a finite number of elements, exemplified by set S above, is called a *finite set*. Set I and set J, each with an infinite number of elements, are, on the other hand, examples of an *infinite set*. Finite sets are always *denumerable* (or *countable*), i.e., their elements can be counted one by one in the sequence $1, 2, 3, \ldots$. Infinite sets may, however, be either denumerable (set I above), or *nondenumerable* (set J above). In the latter case, there is no way to associate the elements of the set with the natural counting numbers $1, 2, 3, \ldots$, and thus the set is not countable.

Membership in a set is indicated by the symbol $\in$ (a variant of the Greek letter epsilon ϵ for "element"), which is read: "is an element of." Thus, for the two sets S and I defined above, we may write

$$2 \in S \qquad 3 \in S \qquad 8 \in I \qquad 9 \in I \qquad \text{(etc.)}$$

but obviously $8 \notin S$ (read: "8 is not an element of set S"). If we use the symbol R to denote the set of all real numbers, then the statement "x is some real number" can be simply expressed by

$$x \in R$$

Relationships between Sets

When two sets are compared with each other, several possible kinds of relationship may be observed. If two sets S_1 and S_2 happen to contain identical elements,

$$S_1 = \{2, 7, a, f\} \qquad \text{and} \qquad S_2 = \{2, a, 7, f\}$$

then S_1 and S_2 are said to be *equal* ($S_1 = S_2$). Note that the order of appearance of the elements in a set is immaterial. Whenever even one element is different, however, two sets are not equal.

Another kind of relationship is that one set may be a *subset* of another set. If we have two sets

$$S = \{1, 3, 5, 7, 9\} \qquad \text{and} \qquad T = \{3, 7\}$$

then T is a subset of S, because every element of T is also an element of S. A more formal statement of this is: T is a subset of S if and only if "$x \in T$" implies "$x \in S$." Using the set inclusion symbols $\subset$ (is contained in) and $\supset$ (includes), we may then write

$$T \subset S \qquad \text{or} \qquad S \supset T$$

It is possible that two given sets happen to be subsets of each other. When this occurs, however, we can be sure that these two sets are equal. To state this formally: we can have $S_1 \subset S_2$ and $S_2 \subset S_1$ if and only if $S_1 = S_2$.

Note that, whereas the $\in$ symbol relates an individual *element* to a *set*, the $\subset$ symbol relates a *subset* to a *set*. As an application of this idea, we may state on the basis of Fig. 2.1 that the set of all integers is a subset of the set of all rational numbers. Similarly, the set of all rational numbers is a subset of the set of all real numbers.

How many subsets can be formed from the five elements in the set $S = \{1, 3, 5, 7, 9\}$? First of all, each individual element of S can count as a distinct subset of S, such as $\{1\}$, $\{3\}$, etc. But so can any pair, triple, or quadruple of these elements, such as $\{1, 3\}$, $\{1, 5\}, \ldots, \{3, 7, 9\}$, etc. For that matter, the set S itself (with all its five elements) can be considered as one of its own subsets—every element of S is an element of S, and thus the set S itself fulfills the definition of a subset. This is, of course, a limiting case, that from which we get the "largest" possible subset of S, namely, S itself.

At the other extreme, the "smallest" possible subset of S is a set that contains no element at all. Such a set is called the *null set*, or *empty set*, denoted by the symbol $\varnothing$ or $\{\ \}$. The reason for considering the null set as a subset of S is quite interesting: If the null set is not a subset of S ($\varnothing \not\subset S$), then $\varnothing$ must contain at least one element x such that $x \notin S$. But since by definition the null set has no element whatsoever, we cannot say that $\varnothing \not\subset S$; hence the null set is a subset of S.

Counting all the subsets of S, including the two limiting cases S and $\varnothing$, we find a total of $2^5 = 32$ subsets. In general, if a set has n elements, a total of 2^n subsets can be formed from those elements.*

It is extremely important to distinguish the symbol $\varnothing$ or $\{\ \}$ clearly from the notation $\{0\}$; the former is devoid of elements, but the latter does contain an element, zero. The null set is unique; there is only one such set in the whole world, and it is considered a subset of *any* set that can be conceived.

As a third possible type of relationship, two sets may have no elements in common at all. In that case, the two sets are said to be *disjoint*. For example, the set of all positive integers and the set of all negative integers are disjoint sets. A fourth type of relationship occurs when two sets have some elements in common but some elements peculiar to each. In that event, the two sets are neither equal nor disjoint; also, neither set is a subset of the other.

Operations on Sets

When we add, subtract, multiply, divide, or take the square root of some numbers, we are performing mathematical operations. Sets are different from

* Given a set with n elements $\{a, b, c, \ldots, n\}$ we may first classify its subsets into two categories: one with the element a in it, and one without. Each of these two can be further classified into two subcategories: one with the element b in it, and one without. Note that by considering the second element b, we double the number of categories in the classification from 2 to 4 ($= 2^2$). By the same token, the consideration of the element c will increase the total number of categories to 8 ($= 2^3$). When all n elements are considered, the total number of categories will become the total number of subsets, and that number is 2^n.

numbers, but one can similarly perform certain mathematical operations on them. Three principal operations to be discussed here involve the union, intersection, and complement of sets.

To take the *union* of two sets A and B means to form a new set containing those elements (and only those elements) belonging to A, or to B, or to both A and B. The union set is symbolized by $A \cup B$ (read: "A union B").

Example 1 If $A = \{3, 5, 7\}$ and $B = \{2, 3, 4, 8\}$, then

$$A \cup B = \{2, 3, 4, 5, 7, 8\}$$

This example illustrates the case in which two sets A and B are neither equal nor disjoint and in which neither is a subset of the other.

Example 2 Again referring to Fig. 2.1, we see that the union of the set of all integers and the set of all fractions is the set of all rational numbers. Similarly, the union of the rational-number set and the irrational-number set yields the set of all real numbers.

The *intersection* of two sets A and B, on the other hand, is a new set which contains those elements (and only those elements) belonging to *both A and B*. The intersection set is symbolized by $A \cap B$ (read: "A intersection B").

Example 3 From the sets A and B in Example 1, we can write

$$A \cap B = \{3\}$$

Example 4 If $A = \{-3, 6, 10\}$ and $B = \{9, 2, 7, 4\}$, then $A \cap B = \varnothing$. Set A and set B are disjoint; therefore their intersection is the empty set—no element is common to A and B.

It is obvious that intersection is a more restrictive concept than union. In the former, only the elements *common to A and B* are acceptable, whereas in the latter, membership in *either A or B* is sufficient to establish membership in the union set. The operator symbols $\cap$ and $\cup$—which, incidentally, have the same kind of general status as the symbols $\sqrt{\ }$, $+$, $\div$, etc.—therefore have the connotations "and" and "or," respectively. This point can be better appreciated by comparing the following formal definitions of intersection and union:

Intersection: $\quad A \cap B = \{x \mid x \in A \text{ and } x \in B\}$

Union: $\quad A \cup B = \{x \mid x \in A \text{ or } x \in B\}$

Before explaining the *complement* of a set, let us first introduce the concept of *universal set*. In a particular context of discussion, if the only numbers used are the set of the first seven positive integers, we may refer to it as the universal set, U. Then, with a given set, say, $A = \{3, 6, 7\}$, we can define another set $\tilde{A}$ (read: "the complement of A") as the set that contains all the numbers in the universal set U which are not in the set A. That is,

$$\tilde{A} = \{x \mid x \in U \text{ and } x \notin A\} = \{1, 2, 4, 5\}$$

Note that, whereas the symbol $\cup$ has the connotation "or" and the symbol $\cap$ means "and," the complement symbol $\sim$ carries the implication of "not."

Example 5 If $U = \{5, 6, 7, 8, 9\}$ and $A = \{5, 6\}$, then $\tilde{A} = \{7, 8, 9\}$.

Example 6 What is the complement of U? Since every object (number) under consideration is included in the universal set, the complement of U must be empty. Thus $\tilde{U} = \varnothing$.

The three types of set operation can be visualized in the three diagrams of Fig. 2.2, known as *Venn diagrams*. In diagram *a*, the points in the upper circle form a set A, and the points in the lower circle form a set B. The union of A and B then consists of the shaded area covering both circles. In diagram *b* are shown the same two sets (circles). Since their intersection should comprise only the points common to both sets, only the (shaded) overlapping portion of the two circles satisfies the definition. In diagram *c*, let the points in the rectangle be the universal set and let A be the set of points in the circle; then the complement set $\tilde{A}$ will be the (shaded) area outside the circle.

Laws of Set Operations

From Fig. 2.2, it may be noted that the shaded area in diagram *a* represents not only $A \cup B$ but also $B \cup A$. Analogously, in diagram *b* the small shaded area is the visual representation not only of $A \cap B$ but also of $B \cap A$. When formalized,

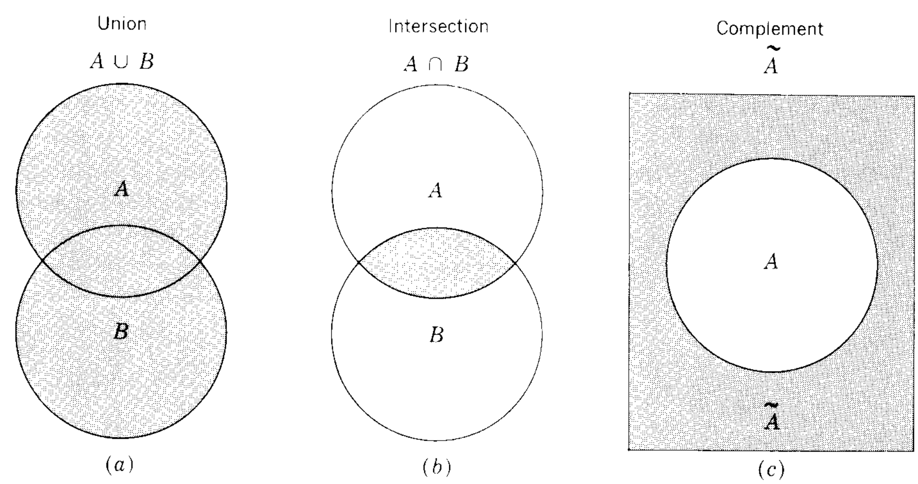

Figure 2.2

this result is known as the *commutative law* (of unions and intersections):

$$A \cup B = B \cup A \qquad A \cap B = B \cap A$$

These relations are very similar to the algebraic laws $a + b = b + a$ and $a \times b = b \times a$.

To take the union of three sets A, B, and C, we first take the union of any two sets and then "union" the resulting set with the third; a similar procedure is applicable to the intersection operation. The results of such operations are illustrated in Fig. 2.3. It is interesting that the order in which the sets are selected for the operation is immaterial. This fact gives rise to the *associative law* (of unions and intersections):

$$A \cup (B \cup C) = (A \cup B) \cup C$$

$$A \cap (B \cap C) = (A \cap B) \cap C$$

These equations are strongly reminiscent of the algebraic laws $a + (b + c) = (a + b) + c$ and $a \times (b \times c) = (a \times b) \times c$.

There is also a law of operation that applies when unions and intersections are used in combination. This is the *distributive law* (of unions and intersections):

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

These resemble the algebraic law $a \times (b + c) = (a \times b) + (a \times c)$.

Example 7 Verify the distributive law, given $A = \{4, 5\}$, $B = \{3, 6, 7\}$, and $C = \{2, 3\}$. To verify the first part of the law, we find the left- and right-hand expressions separately:

Left: $A \cup (B \cap C) = \{4, 5\} \cup \{3\} = \{3, 4, 5\}$

Right: $(A \cup B) \cap (A \cup C) = \{3, 4, 5, 6, 7\} \cap \{2, 3, 4, 5\} = \{3, 4, 5\}$

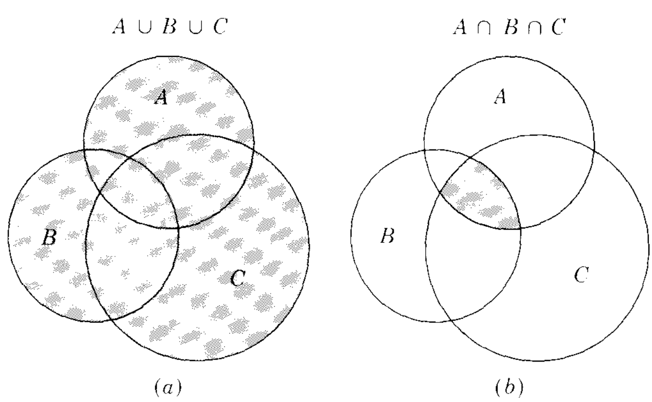

Figure 2.3

Since the two sides yield the same result, the law is verified. Repeating the procedure for the second part of the law, we have

Left: $\quad A \cap (B \cup C) = \{4, 5\} \cap \{2, 3, 6, 7\} = \varnothing$

Right: $\quad (A \cap B) \cup (A \cap C) = \varnothing \cup \varnothing = \varnothing$

Thus the law is again verified.

EXERCISE 2.3

1 Write the following in set notation:
 (*a*) The set of all real numbers greater than 27.
 (*b*) The set of all real numbers greater than 8 but less than 73.

2 Given the sets $S_1 = \{2, 4, 6\}$, $S_2 = \{7, 2, 6\}$, $S_3 = \{4, 2, 6\}$, and $S_4 = \{2, 4\}$, which of the following statements are true?

(*a*) $S_1 = S_2$	(*d*) $3 \notin S_2$	(*g*) $S_1 \supset S_4$
(*b*) $S_1 = R$	(*e*) $4 \notin S_4$	(*h*) $\varnothing \subset S_2$
(*c*) $5 \in S_2$	(*f*) $S_4 \subset R$	(*i*) $S_3 \supset \{1, 2\}$

3 Referring to the four sets given in the preceding problem, find:

(*a*) $S_1 \cup S_2$	(*c*) $S_2 \cap S_3$	(*e*) $S_4 \cap S_2 \cap S_1$
(*b*) $S_1 \cup S_3$	(*d*) $S_2 \cap S_4$	(*f*) $S_3 \cup S_1 \cup S_4$

4 Which of the following statements are valid?

(*a*) $A \cup A = A$	(*e*) $A \cap \varnothing = \varnothing$
(*b*) $A \cap A = A$	(*f*) $A \cap U = A$
(*c*) $A \cup \varnothing = A$	(*g*) The complement of $\tilde{A}$ is A.
(*d*) $A \cup U = U$	

5 Given $A = \{4, 5, 6\}$, $B = \{3, 4, 6, 7\}$, and $C = \{2, 3, 6\}$, verify the distributive law.

6 Verify the distributive law by means of Venn diagrams, with different orders of successive shading.

7 Enumerate all the subsets of the set $\{a, b, c\}$.

8 Enumerate all the subsets of the set $S = \{1, 3, 5, 7\}$. How many subsets are there altogether?

9 Example 6 shows that $\varnothing$ is the complement of U. But since the null set is a subset of *any* set, $\varnothing$ must be a subset of U. Inasmuch as the term "complement of U" implies the notion of being *not in* U, whereas the term "subset of U" implies the notion of being *in* U, it seems paradoxical for $\varnothing$ to be both of these. How do you resolve this paradox?

2.4 RELATIONS AND FUNCTIONS

Our discussion of sets was prompted by the usage of that term in connection with the various kinds of numbers in our number system. However, sets can refer as well to objects other than numbers. In particular, we can speak of sets of "ordered pairs"—to be defined presently—which will lead us to the important concepts of relations and functions.

Ordered Pairs

In writing a set $\{a, b\}$, we do not care about the order in which the elements a and b appear, because by definition $\{a, b\} = \{b, a\}$. The pair of elements a and b is in this case an *unordered pair*. When the ordering of a and b does carry a significance, however, we can write two different *ordered pairs* denoted by (a, b) and (b, a), which have the property that $(a, b) \neq (b, a)$ unless $a = b$. Similar concepts apply to a set with more than two elements, in which case we can distinguish between ordered and unordered triples, quadruples, quintuples, and so forth. Ordered pairs, triples, etc., collectively can be called *ordered sets*.

Example 1 To show the age and weight of each student in a class, we can form ordered pairs (a, w), in which the first element indicates the age (in years) and the second element indicates the weight (in pounds). Then (19,127) and (127,19) would obviously mean different things. Moreover, the latter ordered pair would hardly fit any student anywhere.

Example 2 When we speak of the set of the five finalists in a contest, the order in which they are listed is of no consequence and we have an unordered quintuple. But after they are judged, respectively, as the winner, first runner-up, etc., the list becomes an ordered quintuple.

Ordered pairs, like other objects, can be elements of a set. Consider the rectangular (cartesian) coordinate plane in Fig. 2.4, where an x axis and a y axis cross each other at a right angle, dividing the plane into four quadrants. This xy plane is an infinite set of points, each of which represents an ordered pair whose first element is an x value and the second element a y value. Clearly, the point labeled $(4, 2)$ is different from the point $(2, 4)$; thus ordering is significant here.

With this visual understanding, we are ready to consider the process of generation of ordered pairs. Suppose, from two given sets, $x = \{1, 2\}$ and $y = \{3, 4\}$, we wish to form all the possible ordered pairs with the first element taken from set x and the second element taken from set y. The result will, of course, be the set of four ordered pairs $(1, 3)$, $(1, 4)$, $(2, 3)$, and $(2, 4)$. This set is called the *cartesian product* (named after Descartes), or *direct product*, of the sets x and y and is denoted by $x \times y$ (read: "x cross y"). It is important to remember that, while x and y are sets of numbers, the cartesian product turns out to be a set of ordered pairs. By enumeration, or by description, we may express the cartesian product alternatively as

$$x \times y = \{(1, 3), (1, 4), (2, 3), (2, 4)\}$$

or

$$x \times y = \{(a, b) \mid a \in x \text{ and } b \in y\}$$

The latter expression may in fact be taken as the general definition of cartesian product for any given sets x and y.

To broaden our horizon, now let both x and y include all the real numbers. Then the resulting cartesian product

$$(2.3) \qquad x \times y = \{(a, b) \mid a \in R \text{ and } b \in R\}$$

will represent the set of all ordered pairs with real-valued elements. Besides, each ordered pair corresponds to a *unique* point in the cartesian coordinate plane of Fig. 2.4, and, conversely, each point in the coordinate plane also corresponds to a *unique* ordered pair in the set $x \times y$. In view of this double uniqueness, a *one-to-one correspondence* is said to exist between the set of ordered pairs in the cartesian product (2.3) and the set of points in the rectangular coordinate plane. The rationale for the notation $x \times y$ is now easy to perceive; we may associate it with the crossing of the x axis and the y axis in Fig. 2.4. A simpler way of expressing the set $x \times y$ in (2.3) is to write it directly as $R \times R$; this is also commonly denoted by R^2.

Extending this idea, we may also define the cartesian product of three sets x, y, and z as follows:

$$x \times y \times z = \{(a, b, c) \mid a \in x, b \in y, c \in z\}$$

which is a set of ordered triples. Furthermore, if the sets x, y, and z each consist of all the real numbers, the cartesian product will correspond to the set of all points in a three-dimensional space. This may be denoted by $R \times R \times R$, or

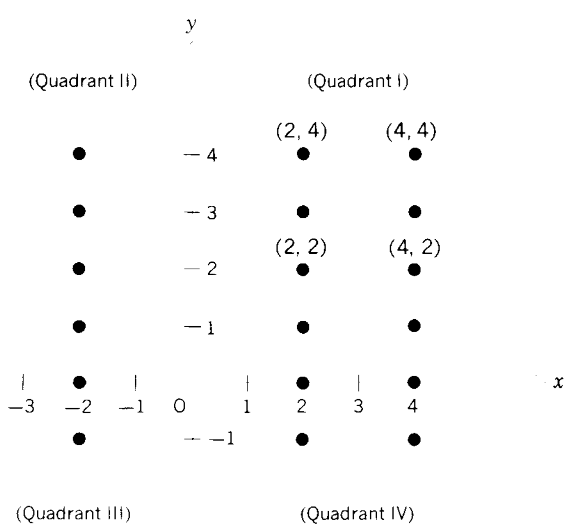

Figure 2.4

more simply, R^3. In the following development, all the variables are taken to be real-valued; thus the framework of our discussion will generally be R^2, or $R^3, \ldots,$ or R^n.

Relations and Functions

Since any ordered pair associates a y value with an x value, any collection of ordered pairs—any subset of the cartesian product (2.3)—will constitute a *relation* between y and x. Given an x value, one or more y values will be specified by that relation. For convenience, we shall now write the elements of $x \times y$ generally as (x, y)—rather than as (a, b), as was done in (2.3)—where both x and y are variables.

Example 3 The set $\{(x, y) \mid y = 2x\}$ is a set of ordered pairs including, for example, $(1, 2)$, $(0, 0)$, and $(-1, -2)$. It constitutes a relation, and its graphical counterpart is the set of points lying on the straight line $y = 2x$, as seen in Fig. 2.5.

Example 4 The set $\{(x, y) \mid y \leq x\}$, which consists of such ordered pairs as $(1, 0)$, $(1, 1)$, and $(1, -4)$, constitutes another relation. In Fig. 2.5, this set corresponds to the set of all points in the shaded area which satisfy the inequality $y \leq x$.

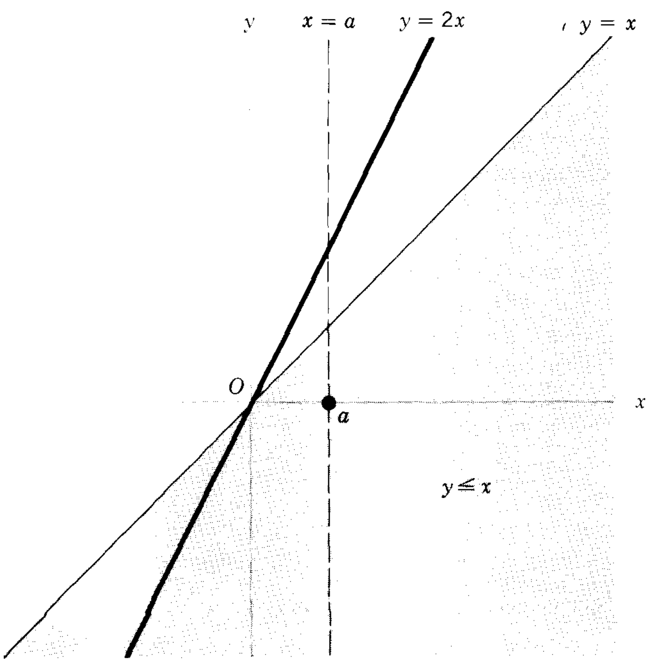

Figure 2.5

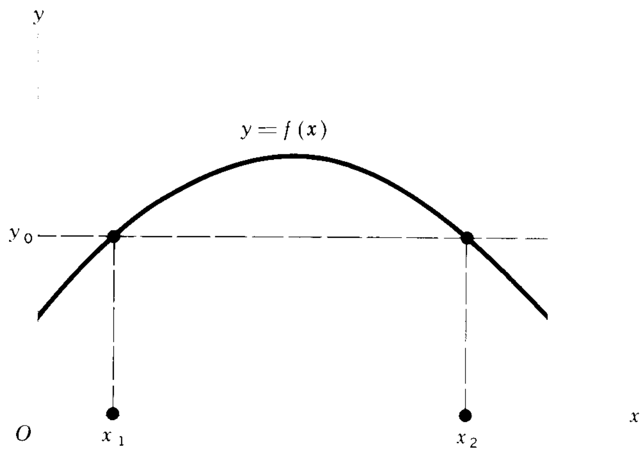

Figure 2.6

Observe that, when the x value is given, it may not always be possible to determine a *unique* y value from a relation. In Example 4, the three exemplary ordered pairs show that if $x = 1$, y can take various values, such as 0, 1, or -4, and yet in each case satisfy the stated relation. Graphically, two or more points of a relation may fall on a single vertical line in the xy plane. This is exemplified in Fig. 2.5, where many points in the shaded area (representing the relation $y \leq x$) fall on the broken vertical line labeled $x = a$.

As a special case, however, a relation may be such that for each x value there exists only *one* corresponding y value. The relation in Example 3 is a case in point. In that case, y is said to be a *function* of x, and this is denoted by $y = f(x)$, which is read: "y equals f of x." [*Note:* $f(x)$ does *not* mean f times x.] A function is therefore a set of ordered pairs with the property that any x value *uniquely* determines a y value.* It should be clear that a function must be a relation, but a relation may not be a function.

Although the definition of a function stipulates a unique y for each x, the converse is not required. In other words, more than one x value may legitimately be associated with the same y value. This possibility is illustrated in Fig. 2.6, where the values x_1 and x_2 in the x set are both associated with the same value (y_0) in the y set by the function $y = f(x)$.

A function is also called a *mapping*, or *transformation*; both words connote the action of associating one thing with another. In the statement $y = f(x)$, the functional notation f may thus be interpreted to mean a rule by which the set x is "mapped" ("transformed") into the set y. Thus we may write

$$f: x \to y$$

* This definition of "function" corresponds to what would be called a *single-valued function* in the older terminology. What was formerly called a *multivalued function* is now referred to as a *relation*.

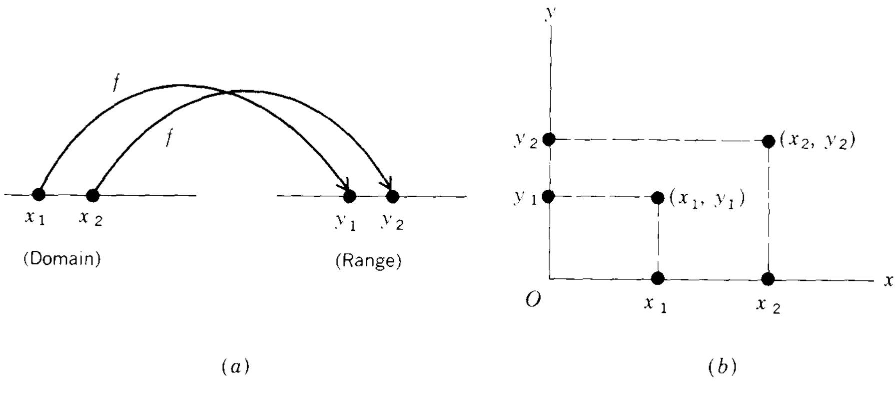

Figure 2.7

where the arrow indicates mapping, and the letter f symbolically specifies a rule of mapping. Since f represents a *particular* rule of mapping, a different functional notation must be employed to denote another function that may appear in the same model. The customary symbols (besides f) used for this purpose are g, F, G, the Greek letters ϕ (phi) and ψ (psi), and their capitals, Φ and Ψ. For instance, two variables y and z may both be functions of x, but if one function is written as $y = f(x)$, the other should be written as $z = g(x)$, or $z = \phi(x)$. It is also permissible, however, to write $y = y(x)$ and $z = z(x)$, thereby dispensing with the symbols f and g entirely.

In the function $y = f(x)$, x is referred to as the *argument* of the function, and y is called the *value* of the function. We shall also alternatively refer to x as the *independent variable* and y as the *dependent variable*. The set of all permissible values that x can take in a given context is known as the *domain* of the function, which may be a subset of the set of all real numbers. The y value into which an x value is mapped is called the *image* of that x value. The set of all images is called the *range* of the function, which is the set of all values that the y variable will take. Thus the domain pertains to the independent variable x, and the range has to do with the dependent variable y.

As illustrated in Fig. 2.7*a*, we may regard the function f as a rule for mapping each point on some line segment (the domain) into some point on another line segment (the range). By placing the domain on the x axis and the range on the y axis, as in diagram *b*, however, we immediately obtain the familiar two-dimensional graph, in which the association between x values and y values is specified by a set of ordered pairs such as (x_1, y_1) and (x_2, y_2).

In economic models, behavioral equations usually enter as functions. Since most variables in economic models are by their nature restricted to being nonnegative real numbers,* their domains are also so restricted. This is why most

* We say "nonnegative" rather than "positive" when zero values are permissible.

geometric representations in economics are drawn only in the first quadrant. In general, we shall not bother to specify the domain of every function in every economic model. When no specification is given, it is to be understood that the domain (and the range) will only include numbers for which a function makes economic sense.

Example 5 The total cost C of a firm per day is a function of its daily output Q: $C = 150 + 7Q$. The firm has a capacity limit of 100 units of output per day. What are the domain and the range of the cost function? Inasmuch as Q can vary only between 0 and 100, the domain is the set of values $0 \le Q \le 100$; or more formally,

$$\text{Domain} = \{Q \mid 0 \le Q \le 100\}$$

As for the range, since the function plots as a straight line, with the minimum C value at 150 (when $Q = 0$) and the maximum C value at 850 (when $Q = 100$), we have

$$\text{Range} = \{C \mid 150 \le C \le 850\}$$

Beware, however, that the extreme values of the range may not always occur where the extreme values of the domain are attained.

EXERCISE 2.4

1 Given $S_1 = \{3, 6, 9\}$, $S_2 = \{a, b\}$, and $S_3 = \{m, n\}$, find the cartesian products:
   (*a*) $S_1 \times S_2$ (*b*) $S_2 \times S_3$ (*c*) $S_3 \times S_1$

2 From the information in the preceding problem, find the cartesian product $S_1 \times S_2 \times S_3$.

3 In general, is it true that $S_1 \times S_2 = S_2 \times S_1$? Under what conditions will these two cartesian products be equal?

4 Does each of the following, drawn in a rectangular coordinate plane, represent a function?
   (*a*) A circle   (*b*) A triangle   (*c*) A rectangle

5 If the domain of the function $y = 5 + 3x$ is the set $\{x \mid 1 \le x \le 4\}$, find the range of the function and express it as a set.

6 For the function $y = -x^2$, if the domain is the set of all nonnegative real numbers, what will its range be?

2.5 TYPES OF FUNCTION

The expression $y = f(x)$ is a general statement to the effect that a mapping is possible, but the actual rule of mapping is not thereby made explicit. Now let us consider several specific types of function, each representing a different rule of mapping.

Constant Functions

A function whose range consists of only one element is called a *constant function*. As an example, we cite the function

$$y = f(x) = 7$$

which is alternatively expressible as $y = 7$ or $f(x) = 7$, whose value stays the same regardless of the value of x. In the coordinate plane, such a function will appear as a horizontal straight line. In national-income models, when investment (I) is exogenously determined, we may have an investment function of the form $I = \$100$ million, or $I = I_0$, which exemplifies the constant function.

Polynomial Functions

The constant function is actually a "degenerate" case of what are known as *polynomial functions*. The word "polynomial" means "multiterm," and a polynomial function of a single variable x has the general form

$$y = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \tag{2.4}$$

in which each term contains a coefficient as well as a nonnegative-integer power of the variable x. (As will be explained later in this section, we can write $x^1 = x$ and $x^0 = 1$ in general; thus the first two terms may be taken to be a_0x^0 and a_1x^1, respectively.) Note that, instead of the symbols $a, b, c, \ldots,$ we have employed the subscripted symbols $a_0, a_1, \ldots, a_n$ for the coefficients. This is motivated by two considerations: (1) we can economize on symbols, since only the letter a is "used up" in this way; and (2) the subscript helps to pinpoint the location of a particular coefficient in the entire equation. For instance, in (2.4), a_2 is the coefficient of x^2, and so forth.

Depending on the value of the integer n (which specifies the highest power of x), we have several subclasses of polynomial function:

Case of $n = 0$:	$y = a_0$	[*constant* function]
Case of $n = 1$:	$y = a_0 + a_1x$	[*linear* function]
Case of $n = 2$:	$y = a_0 + a_1x + a_2x^2$	[*quadratic* function]
Case of $n = 3$:	$y = a_0 + a_1x + a_2x^2 + a_3x^3$	[*cubic* function]

and so forth. The superscript indicators of the powers of x are called *exponents*. The highest power involved, i.e., the value of n, is often called the *degree* of the polynomial function; a quadratic function, for instance, is a second-degree polynomial, and a cubic function is a third-degree polynomial.* The order in which the several terms appear to the right of the equals sign is inconsequential;

* In the several equations just cited, the last coefficient (a_n) is always assumed to be nonzero; otherwise the function would degenerate into a lower-degree polynomial.

they may be arranged in descending order of power instead. Also, even though we have put the symbol y on the left, it is also acceptable to write $f(x)$ in its place.

When plotted in the coordinate plane, a linear function will appear as a straight line, as illustrated in Fig. 2.8*a*. When $x = 0$, the linear function yields $y = a_0$; thus the ordered pair $(0, a_0)$ is on the line. This gives us the so-called "y intercept" (or *vertical intercept*), because it is at this point that the vertical axis intersects the line. The other coefficient, a_1, measures the *slope* (the steepness of incline) of our line. This means that a unit increase in x will result in an increment in y in the amount of a_1. What Fig. 2.8*a* illustrates is the case of $a_1 > 0$, involving a positive slope and thus an upward-sloping line; if $a_1 < 0$, the line will be downward-sloping.

A quadratic function, on the other hand, plots as a *parabola*—roughly, a curve with a single built-in bump or wiggle. The particular illustration in Fig. 2.8*b* implies a negative a_2; in the case of $a_2 > 0$, the curve will "open" the other way, displaying a valley rather than a hill. The graph of a cubic function will, in general, manifest two wiggles, as illustrated in Fig. 2.8*c*. These functions will be used quite frequently in the economic models discussed below.

Rational Functions

A function such as

$$y = \frac{x - 1}{x^2 + 2x + 4}$$

in which y is expressed as a ratio of two polynomials in the variable x, is known as a *rational function* (again, meaning *ratio*-nal). According to this definition, any polynomial function must itself be a rational function, because it can always be expressed as a ratio to 1, which is a constant function.

A special rational function that has interesting applications in economics is the function

$$y = \frac{a}{x} \qquad \text{or} \qquad xy = a$$

which plots as a *rectangular hyperbola*, as in Fig. 2.8*d*. Since the product of the two variables is always a fixed constant in this case, this function may be used to represent that special demand curve—with price P and quantity Q on the two axes—for which the total expenditure PQ is constant at all levels of price. (Such a demand curve is the one with a unitary elasticity at each point on the curve.) Another application is to the average fixed cost (AFC) curve. With AFC on one axis and output Q on the other, the AFC curve must be rectangular-hyperbolic because AFC $\times$ Q(= total fixed cost) is a fixed constant.

The rectangular hyperbola drawn from $xy = a$ never meets the axes, even if extended indefinitely upward and to the right. Rather, the curve approaches the axes *asymptotically*: as y becomes very large, the curve will come ever closer to the

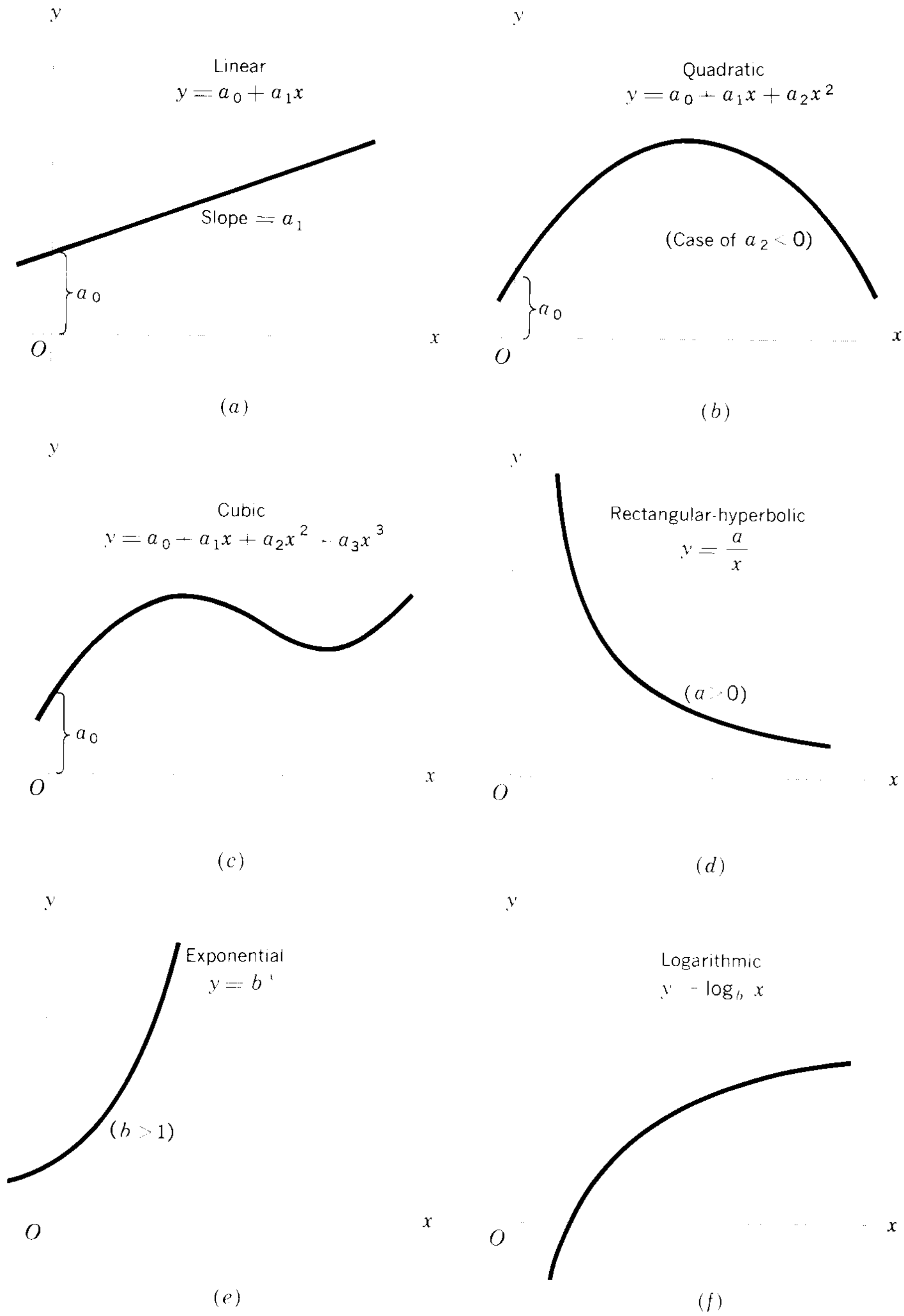
Linear
$y = a_0 + a_1 x$
Slope $= a_1$
a_0
(a)
Quadratic
$y = a_0 + a_1 x + a_2 x^2$
(Case of $a_2 < 0$)
a_0
(b)
Cubic
$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3$
a_0
(c)
Rectangular-hyperbolic
$y = \frac{a}{x}$
($a > 0$)
(d)
Exponential
$y = b^x$
($b > 1$)
(e)
Logarithmic
$y = \log_b x$
(f)

Figure 2.8

y axis but never actually reach it, and similarly for the x axis. The axes constitute the *asymptotes* of this function.

Nonalgebraic Functions

Any function expressed in terms of polynomials and/or roots (such as square root) of polynomials is an *algebraic function*. Accordingly, the functions discussed thus far are all algebraic. A function such as $y = \sqrt{x^2 + 3}$ is not rational, yet it is algebraic.

However, *exponential functions* such as $y = b^x$, in which the independent variable appears in the exponent, are *nonalgebraic*. The closely related *logarithmic functions*, such as $y = \log_b x$, are also nonalgebraic. These two types of function will be explained in detail in Chap. 10, but their general graphic shapes are indicated in Fig. 2.8e and f. Other types of nonalgebraic function are the *trigonometric* (or *circular*) *functions*, which we shall discuss in Chap. 15 in connection with dynamic analysis. We should add here that nonalgebraic functions are also known by the more esoteric name of *transcendental functions*.

A Digression on Exponents

In discussing polynomial functions, we introduced the term *exponents* as indicators of the power to which a variable (or number) is to be raised. The expression 6^2 means that 6 is to be raised to the second power; that is, 6 is to be multiplied by itself, or $6^2 \equiv 6 \times 6 = 36$. In general, we define

$$x^n \equiv \underbrace{x \times x \times \cdots \times x}_{n \text{ terms}}$$

and as a special case, we note that $x^1 = x$. From the general definition, it follows that exponents obey the following rules:

Rule I $\quad x^m \times x^n = x^{m+n} \qquad$ (for example, $x^3 \times x^4 = x^7$)

PROOF

$$x^m \times x^n = \Big(\underbrace{x \times x \times \cdots \times x}_{m \text{ terms}}\Big)\Big(\underbrace{x \times x \times \cdots \times x}_{n \text{ terms}}\Big) = \underbrace{x \times x \times \cdots \times x}_{m+n \text{ terms}} = x^{m+n}$$

Rule II $\quad \dfrac{x^m}{x^n} = x^{m-n} \qquad (x \neq 0) \qquad \left(\text{for example, } \dfrac{x^4}{x^3} = x\right)$

PROOF

$$\frac{x^m}{x^n} = \frac{\overbrace{x \times x \times \cdots \times x}^{m \text{ terms}}}{\underbrace{x \times x \times \cdots \times x}_{n \text{ terms}}} = \underbrace{x \times x \times \cdots \times x}_{m-n \text{ terms}} = x^{m-n}$$

because the n terms in the denominator cancel out n of the m terms in the numerator. Note that the case of $x = 0$ is ruled out in the statement of this rule. This is because when $x = 0$, the expression x^m/x^n would involve division by zero, which is undefined.

What if $m < n$: say, $m = 2$ and $n = 5$? In that case we get, according to Rule II, $x^{m-n} = x^{-3}$, a *negative power* of x. What does this mean? The answer is actually supplied by Rule II itself: When $m = 2$ and $n = 5$, we have

$$\frac{x^2}{x^5} = \frac{x \times x}{x \times x \times x \times x \times x} = \frac{1}{x \times x \times x} = \frac{1}{x^3}$$

Thus $x^{-3} = 1/x^3$, and this may be generalized into another rule:

Rule III $\quad x^{-n} = \dfrac{1}{x^n} \qquad (x \neq 0)$

To raise a (nonzero) number to a power of *minus n* is to take the *reciprocal* of its nth power.

Another special case in the application of Rule II is when $m = n$, which yields the expression $x^{m-n} = x^{m-m} = x^0$. To interpret the meaning of raising a number x to the zeroth power, we can write out the term x^{m-m} in accordance with Rule II above, with the result that $x^m/x^m = 1$. Thus we may conclude that any (nonzero) number raised to the zeroth power is equal to 1. (The expression 0^0 is undefined.) This may be expressed as another rule:

Rule IV $\quad x^0 = 1 \qquad (x \neq 0)$

As long as we are concerned only with polynomial functions, only (nonnegative) integer powers are required. In exponential functions, however, the exponent is a variable that can take noninteger values as well. In order to interpret a number such as $x^{1/2}$, let us consider the fact that, by Rule I above, we have

$$x^{1/2} \times x^{1/2} = x^1 = x$$

Since $x^{1/2}$ multiplied by itself is x, $x^{1/2}$ must be the square root of x. Similarly, $x^{1/3}$ can be shown to be the cube root of x. In general, therefore, we can state the following rule:

Rule V $\quad x^{1/n} = \sqrt[n]{x}$

Two other rules obeyed by exponents are:

Rule VI $\quad (x^m)^n = x^{mn}$

Rule VII $\quad x^m \times y^m = (xy)^m$

EXERCISE 2.5

1 Graph the functions

(a) $y = 8 + 3x$ (b) $y = 8 - 3x$ (c) $y = 3x + 12$

(In each case, consider the domain as consisting of nonnegative real numbers only.)

2 What is the major difference between (a) and (b) above? How is this difference reflected in the graphs? What is the major difference between (a) and (c)? How do their graphs reflect it?

3 Graph the functions

(a) $y = -x^2 + 5x - 2$ (b) $y = x^2 + 5x - 2$

with the set of values $-5 \le x \le 5$ as the domain. It is well known that the sign of the coefficient of the x^2 term determines whether the graph of a quadratic function will have a "hill" or a "valley." On the basis of the present problem, which sign is associated with the hill? Supply an intuitive explanation for this.

4 Graph the function $y = 36/x$, assuming that x and y can take positive values only. Next, suppose that both variables can take negative values as well; how must the graph be modified to reflect this change in assumption?

5 Condense the following expressions:

(a) $x^4 \times x^{15}$ (b) $x^a \times x^b \times x^c$ (c) $x^3 \times y^3 \times z^3$

6 Find: (a) x^3/x^{-3} (b) $(x^{1/2} \times x^{1/3})/x^{2/3}$

7 Show that $x^{m/n} = \sqrt[n]{x^m} = \left(\sqrt[n]{x}\right)^m$. Specify the rules applied in each step.

8 Prove Rule VI and Rule VII.

2.6 FUNCTIONS OF TWO OR MORE INDEPENDENT VARIABLES

Thus far, we have considered only functions of a single independent variable, $y = f(x)$. But the concept of a function can be readily extended to the case of two or more independent variables. Given a function

$$z = g(x, y)$$

a given pair of x and y values will uniquely determine a value of the dependent variable z. Such a function is exemplified by

$$z = ax + by \qquad \text{or} \qquad z = a_0 + a_1x + a_2x^2 + b_1y + b_2y^2$$

Just as the function $y = f(x)$ maps a point in the domain into a point in the range, the function g will do precisely the same. However, the domain is in this case no longer a set of numbers but a set of ordered pairs (x, y), because we can determine z only when *both* x *and* y are specified. The function g is thus a mapping from a point in a two-dimensional space into a point on a line segment (i.e., a point in a one-dimensional space), such as from the point (x_1, y_1) into the point z_1 or from (x_2, y_2) into z_2 in Fig. 2.9*a*.

If a vertical z axis is erected perpendicular to the xy plane, as is done in diagram *b*, however, there will result a three-dimensional space in which the function g can be given a graphical representation as follows. The domain of the function will be some subset of the points in the xy plane, and the value of the function (value of z) for a given point in the domain—say, (x_1, y_1)—can be indicated by the height of a vertical line planted on that point. The association between the three variables is thus summarized by the ordered triple (x_1, y_1, z_1), which is a specific point in the three-dimensional space. The locus of such ordered triples, which will take the form of a *surface*, then constitutes the graph of the function g. Whereas the function $y = f(x)$ is a set of ordered *pairs*, the function

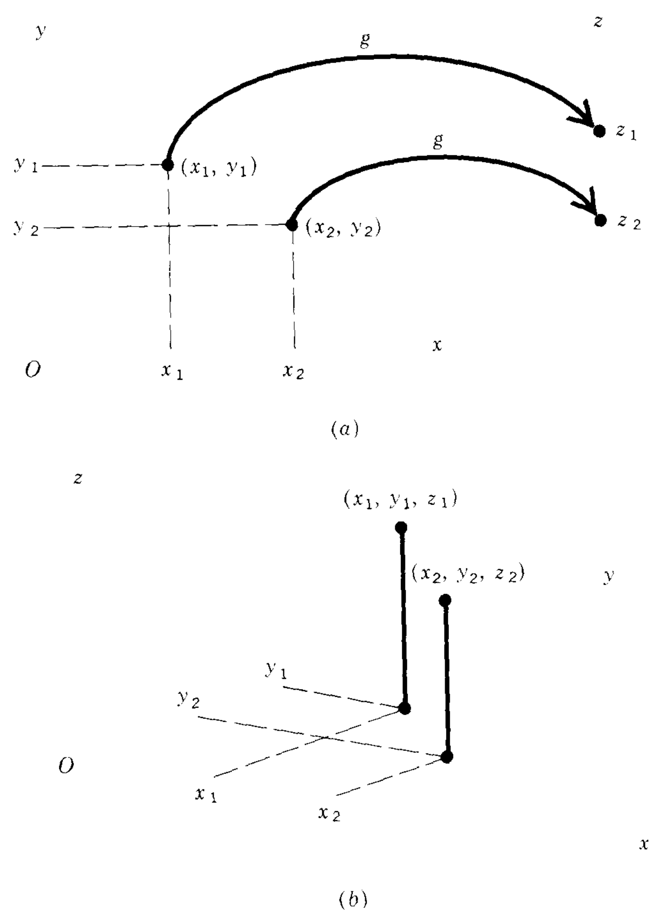

Figure 2.9

$z = g(x, y)$ will be a set of ordered *triples*. We shall have many occasions to use functions of this type in economic models. One ready application is in the area of production functions. Suppose that output is determined by the amounts of capital (K) and labor (L) employed; then we can write a production function in the general form $Q = Q(K, L)$.

The possibility of further extension to the cases of three or more independent variables is now self-evident. With the function $y = h(u, v, w)$, for example, we can map a point in the three-dimensional space, (u_1, v_1, w_1), into a point in a one-dimensional space (y_1). Such a function might be used to indicate that a consumer's utility is a function of his consumption of three different commodities, and the mapping is from a three-dimensional commodity space into a one-dimensional utility space. But this time it will be physically impossible to graph the function, because for that task a four-dimensional diagram is needed to picture the ordered quadruples, but the world in which we live is only three-dimensional. Nonetheless, in view of the intuitive appeal of geometric analogy, we can continue to refer to an ordered quadruple (u_1, v_1, w_1, y_1) as a "point" in the four-dimensional space. The locus of such points will give the (nongraphable) graph of the function $y = h(u, v, w)$, which is called a *hypersurface*. These terms, viz., point and hypersurface, are also carried over to the general case of the n-dimensional space.

Functions of more than one variable can be classified into various types, too. For instance, a function of the form

$$y = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

is a *linear* function, whose characteristic is that every variable is raised to the first power only. A *quadratic* function, on the other hand, involves first and second powers of one or more independent variables, but the sum of exponents of the variables appearing in any single term must not exceed two.

Note that instead of denoting the independent variables by x, u, v, w, etc., we have switched to the symbols $x_1, x_2, \ldots, x_n$. The latter notation, like the system of subscripted coefficients, has the merit of economy of alphabet, as well as of an easier accounting of the number of variables involved in a function.

2.7 LEVELS OF GENERALITY

In discussing the various types of function, we have without explicit notice introduced examples of functions that pertain to varying levels of generality. In certain instances, we have written functions in the form

$$y = 7 \qquad y = 6x + 4 \qquad y = x^2 - 3x + 1 \qquad \text{(etc.)}$$

Not only are these expressed in terms of numerical coefficients, but they also indicate specifically whether each function is constant, linear, or quadratic. In terms of graphs, each such function will give rise to a well-defined unique curve. In view of the numerical nature of these functions, the solutions of the model based on them will emerge as numerical values also. The drawback is that, if we wish to know how our analytical conclusion will change when a different set of numerical coefficients comes into effect, we must go through the reasoning process afresh each time. Thus, the results obtained from specific functions have very little generality.

On a more general level of discussion and analysis, there are functions in the form

$$y = a \qquad y = a + bx \qquad y = a + bx + cx^2 \qquad \text{(etc.)}$$

Since parameters are used, each function represents not a single curve but a whole family of curves. The function $y = a$, for instance, encompasses not only the specific cases $y = 0$, $y = 1$, and $y = 2$ but also $y = \frac{1}{3}$, $y = -5, \ldots,$ ad infinitum. With parametric functions, the outcome of mathematical operations will also be in terms of parameters. These results are more general in the sense that, by assigning various values to the parameters appearing in the solution of the model, a whole family of specific answers may be obtained without having to repeat the reasoning process anew.

In order to attain an even higher level of generality, we may resort to the general function statement $y = f(x)$, or $z = g(x, y)$. When expressed in this form, the function is not restricted to being either linear, quadratic, exponential, or trigonometric—all of which are subsumed under the notation. The analytical result based on such a general formulation will therefore have the most general applicability. As will be found below, however, in order to obtain economically meaningful results, it is often necessary to impose certain qualitative restrictions on the general functions built into a model, such as the restriction that a demand function have a negatively sloped graph or that a consumption function have a graph with a positive slope of less than 1.

To sum up the present chapter, the structure of a mathematical economic model is now clear. In general, it will consist of a system of equations, which may be definitional, behavioral, or in the nature of equilibrium conditions.* The behavioral equations are usually in the form of functions, which may be linear or nonlinear, numerical or parametric, and with one independent variable or many. It is through these that the analytical assumptions adopted in the model are given mathematical expression.

In attacking an analytical problem, therefore, the first step is to select the appropriate variables—exogenous as well as endogenous—for inclusion in the model. Next, we must translate into equations the set of chosen analytical assumptions regarding the human, institutional, technological, legal, and other behavioral aspects of the environment affecting the working of the variables. Only then can an attempt be made to derive a set of conclusions through relevant mathematical operations and manipulations and to give them appropriate economic interpretations.

* Inequalities may also enter as an important ingredient of a model, but we shall not worry about them for the time being.

PART

TWO

STATIC (OR EQUILIBRIUM) ANALYSIS

CHAPTER

THREE

EQUILIBRIUM ANALYSIS IN ECONOMICS

The analytical procedure outlined in the preceding chapter will first be applied to what is known as *static analysis*, or *equilibrium analysis*. For this purpose, it is imperative first to have a clear understanding of what "equilibrium" means.

3.1 THE MEANING OF EQUILIBRIUM

Like any economic term, *equilibrium* can be defined in various ways. According to one definition, an equilibrium is "a constellation of selected interrelated variables so adjusted to one another that no inherent tendency to change prevails in the model which they constitute."* Several words in this definition deserve special attention. First, the word "selected" underscores the fact that there do exist variables which, by the analyst's choice, have not been included in the model. Hence the equilibrium under discussion can have relevance only in the context of the particular set of variables chosen, and if the model is enlarged to include additional variables, the equilibrium state pertaining to the smaller model will no longer apply.

Second, the word "interrelated" suggests that, in order for equilibrium to obtain, all variables in the model must simultaneously be in a state of rest. Moreover, the state of rest of each variable must be compatible with that of every

* Fritz Machlup, "Equilibrium and Disequilibrium: Misplaced Concreteness and Disguised Politics," *Economic Journal*, March 1958, p. 9. (Reprinted in F. Machlup, *Essays on Economic Semantics*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.)

other variable; otherwise some variable(s) will be changing, thereby also causing the others to change in a chain reaction, and no equilibrium can be said to exist.

Third, the word "inherent" implies that, in defining an equilibrium, the state of rest involved is based only on the balancing of the internal forces of the model, while the external factors are assumed fixed. Operationally, this means that parameters and exogenous variables are treated as constants. When the external factors do actually change, there may result a new equilibrium defined on the basis of the new parameter values, but in defining the new equilibrium, the new parameter values are again assumed to persist and stay unchanged.

In essence, an equilibrium for a specified model is a situation that is characterized by a lack of tendency to change. It is for this reason that the analysis of equilibrium (more specifically, the study of what the equilibrium state is like) is referred to as *statics*. The fact that an equilibrium implies no tendency to change may tempt one to conclude that an equilibrium necessarily constitutes a desirable or ideal state of affairs, on the ground that only in the ideal state would there be a lack of motivation for change. Such a conclusion is unwarranted. Even though a certain equilibrium position may represent a desirable state and something to be striven for—such as a profit-maximizing situation, from the firm's point of view—another equilibrium position may be quite undesirable and therefore something to be avoided, such as an underemployment equilibrium level of national income. The only warranted interpretation is that an equilibrium is a situation which, if attained, would tend to perpetuate itself, barring any changes in the external forces.

The desirable variety of equilibrium, which we shall refer to as *goal equilibrium*, will be treated later in Parts 4 and 6 as optimization problems. In the present chapter, the discussion will be confined to the *nongoal* type of equilibrium, resulting not from any conscious aiming at a particular objective but from an impersonal or suprapersonal process of interaction and adjustment of economic forces. Examples of this are the equilibrium attained by a market under given demand and supply conditions and the equilibrium of national income under given conditions of consumption and investment patterns.

3.2 PARTIAL MARKET EQUILIBRIUM—A LINEAR MODEL

In a static-equilibrium model, the standard problem is that of finding the set of values of the endogenous variables which will satisfy the equilibrium condition of the model. This is because once we have identified those values, we have in effect identified the equilibrium state. Let us illustrate with a so-called "partial-equilibrium market model," i.e., a model of price determination in an isolated market.

Constructing the Model

Since only one commodity is being considered, it is necessary to include only three variables in the model: the quantity demanded of the commodity (Q_d), the quantity supplied of the commodity (Q_s), and its price (P). The quantity is measured, say, in pounds per week, and the price in dollars. Having chosen the variables, our next order of business is to make certain assumptions regarding the working of the market. First, we must specify an equilibrium condition—something indispensable in an equilibrium model. The standard assumption is that equilibrium obtains in the market if and only if the excess demand is zero ($Q_d - Q_s = 0$), that is, if and only if the market is cleared. But this immediately raises the question of how Q_d and Q_s themselves are determined. To answer this, we assume that Q_d is a decreasing linear function of P (as P increases, Q_d decreases). On the other hand, Q_s is postulated to be an increasing linear function of P (as P increases, so does Q_s), with the proviso that no quantity is supplied unless the price exceeds a particular positive level. In all, then, the model will contain one equilibrium condition plus two behavioral equations which govern the demand and supply sides of the market, respectively.

Translated into mathematical statements, the model can be written as:

$$
\begin{aligned}
& Q_d = Q_s \\
(3.1) \qquad & Q_d = a - bP \qquad (a, b > 0) \\
& Q_s = -c + dP \qquad (c, d > 0)
\end{aligned}
$$

Four parameters, a, b, c, and d, appear in the two linear functions, and all of them are specified to be positive. When the demand function is graphed, as in Fig. 3.1, its vertical intercept is at a and its slope is $-b$, which is negative, as required. The supply function also has the required type of slope, d being positive, but its

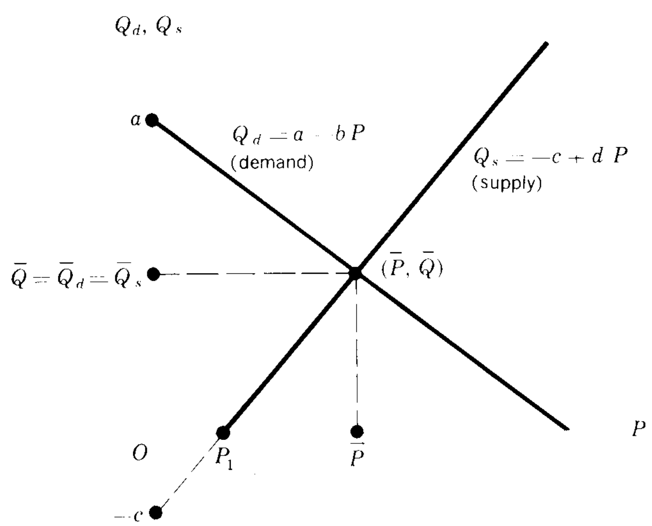

Figure 3.1

vertical intercept is seen to be negative, at $-c$. Why did we want to specify such a negative vertical intercept? The answer is that, in so doing, we force the supply curve to have a positive horizontal intercept at P_1, thereby satisfying the proviso stated earlier that supply will not be forthcoming unless the price is positive and sufficiently high.

The reader should observe that, contrary to the usual practice, quantity rather than price has been plotted vertically in Fig. 3.1. This, however, is in line with the mathematical convention of placing the *dependent* variable on the vertical axis. In a different context below, in which the demand curve is viewed from the standpoint of a business firm as describing the average-revenue curve, $AR \equiv P = f(Q_d)$, we shall reverse the axes and plot P vertically.

With the model thus constructed, the next step is to solve it, i.e., to obtain the solution values of the three endogenous variables, Q_d, Q_s, and P. The solution values, to be denoted $\bar{Q}_d$, $\bar{Q}_s$, and $\bar{P}$, are those values that satisfy the three equations in (3.1) simultaneously; i.e., they are the values which, when substituted into the three equations, make the latter a set of true statements. In the context of an equilibrium model, those values may also be referred to as the *equilibrium values* of the said variables. Since $\bar{Q}_d = \bar{Q}_s$, however, they can be replaced by a single symbol $\bar{Q}$. Hence, an equilibrium solution of the model may simply be denoted by an ordered pair $(\bar{P}, \bar{Q})$. In case the solution is not unique, several ordered pairs may each satisfy the system of simultaneous equations; there will then be a solution set with more than one element in it. However, the multiple-equilibrium situation cannot arise in a linear model such as the present one.

Solution by Elimination of Variables

One way of finding a solution to an equation system is by successive elimination of variables and equations through substitution. In (3.1), the model contains three equations in three variables. However, in view of the equating of Q_d and Q_s by the equilibrium condition, we can let $Q = Q_d = Q_s$ and rewrite the model equivalently as follows:

$$(3.2) \qquad \begin{aligned} Q &= a - bP \\ Q &= -c + dP \end{aligned}$$

thereby reducing the model to two equations in two variables. Moreover, by substituting the first equation into the second in (3.2), the model can be further reduced to a single equation in a single variable:

$$a - bP = -c + dP$$

or, after subtracting $(a + dP)$ from both sides of the equation and multiplying through by -1,

$$(3.3) \qquad (b + d)P = a + c$$

This result is also obtainable directly from (3.1) by substituting the second and third equations into the first.

Since $b + d \neq 0$, it is permissible to divide both sides of (3.3) by $(b + d)$. The result is the solution value of P:

$$\bar{P} = \frac{a + c}{b + d} \tag{3.4}$$

Note that $\bar{P}$ is—as all solution values should be—expressed entirely in terms of the parameters, which represent given data for the model. Thus $\bar{P}$ is a determinate value, as it ought to be. Also note that $\bar{P}$ is positive—as a price should be—because all the four parameters are positive by model specification.

To find the equilibrium quantity $\bar{Q}$ $(= \bar{Q}_d = \bar{Q}_s)$ that corresponds to the value $\bar{P}$, simply substitute (3.4) into *either* equation of (3.2), and then solve the resulting equation. Substituting (3.4) into the demand function, for instance, we can get

$$\bar{Q} = a - \frac{b(a + c)}{b + d} = \frac{a(b + d) - b(a + c)}{b + d} = \frac{ad - bc}{b + d} \tag{3.5}$$

which is again an expression in terms of parameters only. Since the denominator $(b + d)$ is positive, the positivity of $\bar{Q}$ requires that the numerator $(ad - bc)$ be positive as well. Hence, to be economically meaningful, the present model should contain the additional restriction that $ad > bc$.

The meaning of this restriction can be seen in Fig. 3.1. It is well known that the $\bar{P}$ and $\bar{Q}$ of a market model may be determined graphically at the intersection of the demand and supply curves. To have $\bar{Q} > 0$ is to require the intersection point to be located above the horizontal axis in Fig. 3.1, which in turn requires the slopes and vertical intercepts of the two curves to fulfill a certain restriction on their relative magnitudes. That restriction, according to (3.5), is $ad > bc$, given that both b and d are positive.

The intersection of the demand and supply curves in Fig. 3.1, incidentally, is in concept no different from the intersection shown in the Venn diagram of Fig. 2.2*b*. There is one difference only: instead of the points lying within two circles, the present case involves the points that lie on two lines. Let the set of points on the demand and supply curves be denoted, respectively, by D and S. Then, by utilizing the symbol Q $(= Q_d = Q_s)$, the two sets and their intersection can be written

$$D = \{(P, Q) \mid Q = a - bP\}$$

$$S = \{(P, Q) \mid Q = -c + dP\}$$

and $$D \cap S = (\bar{P}, \bar{Q})$$

The intersection set contains in this instance only a single element, the ordered pair $(\bar{P}, \bar{Q})$. The market equilibrium is unique.

EXERCISE 3.2

1 Given the market model

$$Q_d = Q_s$$
$$Q_d = 24 - 2P$$
$$Q_s = -5 + 7P$$

find $\bar{P}$ and $\bar{Q}$ by (*a*) elimination of variables and (*b*) using formulas (3.4) and (3.5). (Use fractions rather than decimals.)

2 Let the demand and supply functions be as follows:

(*a*) $Q_d = 51 - 3P$ (*b*) $Q_d = 30 - 2P$
$Q_s = 6P - 10$ $Q_s = -6 + 5P$

find $\bar{P}$ and $\bar{Q}$ by elimination of variables. (Use fractions rather than decimals.)

3 According to (3.5), for $\bar{Q}$ to be positive, it is necessary that the expression $(ad - bc)$ have the same algebraic sign as $(b + d)$. Verify that this condition is indeed satisfied in the models of the preceding two problems.

4 If $(b + d) = 0$ in the linear market model, can an equilibrium solution be found by using (3.4) and (3.5)? Why or why not?

5 If $(b + d) = 0$ in the linear market model, what can you conclude regarding the positions of the demand and supply curves in Fig. 3.1? What can you conclude, then, regarding the equilibrium solution?

3.3 PARTIAL MARKET EQUILIBRIUM—A NONLINEAR MODEL

Let the linear demand in the isolated market model be replaced by a quadratic demand function, while the supply function remains linear. Then, if numerical coefficients are employed rather than parameters, a model such as the following may emerge:

$$\begin{aligned} Q_d &= Q_s \\ Q_d &= 4 - P^2 \\ Q_s &= 4P - 1 \end{aligned} \tag{3.6}$$

As previously, this system of three equations can be reduced to a single equation by elimination of variables (by substitution):

$$4 - P^2 = 4P - 1$$

or

$$P^2 + 4P - 5 = 0 \tag{3.7}$$

This is a quadratic equation because the left-hand expression is a quadratic function of variable P. The major difference between a quadratic equation and a linear one is that, in general, the former will yield two solution values.

Quadratic Equation versus Quadratic Function

Before discussing the method of solution, a clear distinction should be made between the two terms *quadratic equation* and *quadratic function*. According to the earlier discussion, the expression $P^2 + 4P - 5$ constitutes a quadratic *function*, say, $f(P)$. Hence we may write

(3.8) $$f(P) = P^2 + 4P - 5$$

What (3.8) does is to specify a rule of mapping from P to $f(P)$, such as

P	···	−6	−5	−4	−3	−2	−1	0	1	2	···
$f(P)$	···	7	0	−5	−8	−9	−8	−5	0	7	···

Although we have listed only nine P values in this table, actually *all* the P values in the domain of the function are eligible for listing. It is perhaps for this reason that we rarely speak of "solving" the equation $f(P) = P^2 + 4P - 5$, because we normally expect "solution values" to be few in number, but here all P values can get involved. Nevertheless, one may legitimately consider each ordered pair in the table above—such as $(-6, 7)$ and $(-5, 0)$—as a solution of (3.8), since each such ordered pair indeed satisfies that equation. Inasmuch as an infinite number of such ordered pairs can be written, one for each P value, there is an infinite number of solutions to (3.8). When plotted as a curve, these ordered pairs together yield the parabola in Fig. 3.2.

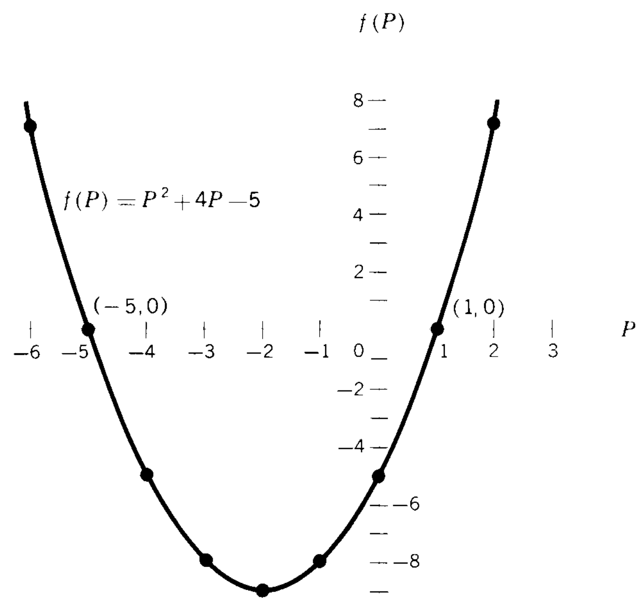

Figure 3.2

In (3.7), where we set the quadratic function $f(P)$ equal to zero, the situation is fundamentally changed. Since the variable $f(P)$ now disappears (having been assigned a zero value), the result is a quadratic *equation* in the single variable P.* Now that $f(P)$ is restricted to a zero value, only a select number of P values can satisfy (3.7) and qualify as its solution values, namely, those P values at which the parabola in Fig. 3.2 intersects the horizontal axis—on which $f(P)$ is zero. Note that this time the solution values are just P values, not ordered pairs. The solution P values are often referred to as the *roots* of the quadratic *equation* $f(P) = 0$, or, alternatively, as the *zeros* of the quadratic *function* $f(P)$.

There are two such intersection points in Fig. 3.2, namely, $(1, 0)$ and $(-5, 0)$. As required, the second element of each of these ordered pairs (the *ordinate* of the corresponding point) shows $f(P) = 0$ in both cases. The first element of each ordered pair (the *abscissa* of the point), on the other hand, gives the solution value of P. Here we get two solutions,

$$\bar{P}_1 = 1 \qquad \text{and} \qquad \bar{P}_2 = -5$$

but only the first is economically admissible, as negative prices are ruled out.

The Quadratic Formula

Equation (3.7) has been solved graphically, but an algebraic method is also available. In general, given a quadratic equation in the form

$$(3.9) \qquad ax^2 + bx + c = 0 \qquad (a \neq 0)$$

its two roots can be obtained from the *quadratic formula*:

$$(3.10) \qquad \bar{x}_1, \bar{x}_2 = \frac{-b \pm (b^2 - 4ac)^{1/2}}{2a}$$

where the $+$ part of the $\pm$ sign yields $\bar{x}_1$ and the $-$ part yields $\bar{x}_2$.

This widely used formula is derived by means of a process known as "completing the square." First, dividing each term of (3.9) by a results in the equation

$$x^2 + \frac{b}{a}x + \frac{c}{a} = 0$$

Subtracting c/a from, and adding $b^2/4a^2$ to, both sides of the equation, we get

$$x^2 + \frac{b}{a}x + \frac{b^2}{4a^2} = \frac{b^2}{4a^2} - \frac{c}{a}$$

* The distinction between quadratic function and quadratic equation just discussed can be extended also to cases of polynomials other than quadratic. Thus, a cubic equation results when a cubic function is set equal to zero.

The left side is now a "perfect square," and thus the equation can be expressed as

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}$$

or, after taking the square root on both sides,

$$x + \frac{b}{2a} = \pm \frac{(b^2 - 4ac)^{1/2}}{2a}$$

Finally, by subtracting $b/2a$ from both sides, the result in (3.10) is evolved.

Applying the formula to (3.7), where $a = 1$, $b = 4$, $c = -5$, and $x = P$, the roots are found to be

$$\bar{P}_1, \bar{P}_2 = \frac{-4 \pm (16 + 20)^{1/2}}{2} = \frac{-4 \pm 6}{2} = 1, -5$$

which check with the graphical solutions in Fig. 3.2. Again, we reject $\bar{P}_2 = -5$ on economic grounds and, after omitting the subscript 1, write simply $\bar{P} = 1$.

With this information in hand, the equilibrium quantity $\bar{Q}$ can readily be found from either the second or the third equation of (3.6) to be $\bar{Q} = 3$.

Another Graphical Solution

One method of graphical solution of the present model has been presented in Fig. 3.2. However, since the quantity variable has been eliminated in deriving the quadratic equation, only $\bar{P}$ can be found from that figure. If we are interested in

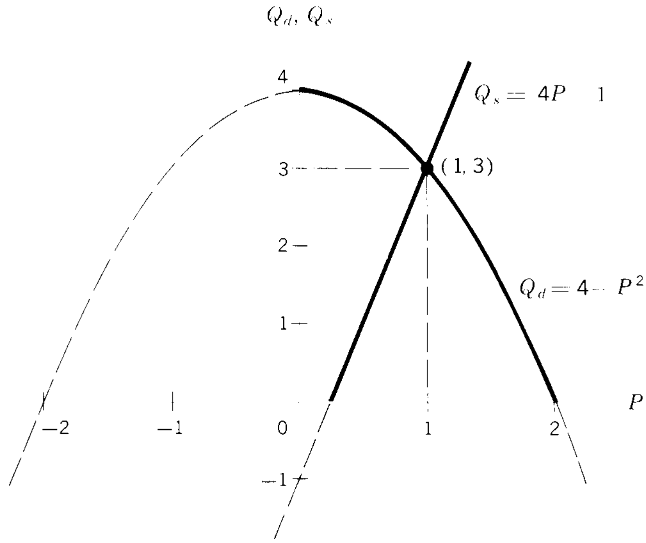

Figure 3.3

finding $\bar{P}$ and $\bar{Q}$ simultaneously from a graph, we must instead use a diagram with Q on one axis and P on the other, similar in construction to Fig. 3.1. This is illustrated in Fig. 3.3. Our problem is of course again to find the intersection of two sets of points, namely,

$$D = \{(P, Q) \mid Q = 4 - P^2\}$$

and $$S = \{(P, Q) \mid Q = 4P - 1\}$$

If no restriction is placed on the domain and the range, the intersection set will contain two elements, namely,

$$D \cap S = \{(1, 3), (-5, -21)\}$$

The former is located in quadrant I, and the latter (not drawn) in quadrant III. If the domain and range are restricted to being nonnegative, however, only the first ordered pair $(1, 3)$ can be accepted. Then the equilibrium is again unique.

Higher-Degree Polynomial Equations

If a system of simultaneous equations reduces not to a linear equation such as (3.3)* or to a quadratic equation such as (3.7) but to a cubic (third-degree polynomial) equation or quartic (fourth-degree polynomial) equation, the roots will be more difficult to find. One useful method which may work is that of *factoring* the function. For example, the expression $x^3 - x^2 - 4x + 4$ can be written as the product of three factors $(x - 1)$, $(x + 2)$, and $(x - 2)$. Thus the cubic equation

$$x^3 - x^2 - 4x + 4 = 0$$

can be written after factoring as

$$(x - 1)(x + 2)(x - 2) = 0$$

In order for the left-hand product to be zero, at least one of the three terms in the product must be zero. Setting each term equal to zero in turn, we get

$$x - 1 = 0 \qquad \text{or} \qquad x + 2 = 0 \qquad \text{or} \qquad x - 2 = 0$$

These three equations will supply the three roots of the cubic equation, namely,

$$\bar{x}_1 = 1 \qquad \bar{x}_2 = -2 \qquad \text{and} \qquad \bar{x}_3 = 2$$

The trick is, of course, to discover the appropriate way of factoring. Unfortunately, no general rule exists, and it must therefore remain a matter of trial and error. Generally speaking, however, given an nth-degree polynomial equation $f(x) = 0$, we can expect exactly n roots, which may be found as follows. First, try to find a constant c_1 such that $f(x)$ is divisible by $(x + c_1)$. The quotient $f(x)/(x + c_1)$ will be a polynomial function of a lesser—$(n - 1)$st—degree; let

* Equation (3.3) can be viewed as the result of setting the linear function $(b + d)P - (a + c)$ equal to zero.

us call it $g(x)$. It then follows that

$$f(x) = (x + c_1)g(x)$$

Now, try to find a constant c_2 such that $g(x)$ is divisible by $(x + c_2)$. The quotient $g(x)/(x + c_2)$ will again be a polynomial function of a lesser—this time $(n - 2)$nd—degree, say, $h(x)$. Since $g(x) = (x + c_2)h(x)$, it follows that

$$f(x) = (x + c_1)g(x) = (x + c_1)(x + c_2)h(x)$$

By repeating the process, it will be possible to reduce the original nth-degree polynomial $f(x)$ to a product of exactly n terms:

$$f(x) = (x + c_1)(x + c_2) \cdots (x + c_n)$$

which, when set equal to zero, will yield n roots. Setting the first factor equal to zero, for example, one gets $\bar{x}_1 = -c_1$. Similarly, the other factors will yield $\bar{x}_2 = -c_2$, $\bar{x}_3 = -c_3$, etc. These results can be more succinctly expressed by employing an *index subscript* i:

$$\bar{x}_i = -c_i \qquad (i = 1, 2, \ldots, n)$$

Even though only one equation is written, the fact that the subscript i can take n different values means that in all there are n equations involved. Thus the index subscript provides a very concise way of statement.

EXERCISE 3.3

1 Find the zeros of the following functions graphically:
(a) $f(x) = x^2 - 7x + 10$ $\qquad$ (b) $g(x) = 2x^2 - 4x - 16$

2 Solve the preceding problem by the quadratic formula.

3 Solve the following polynomial equations by factoring:
(a) $P^2 + 4P - 5 = 0$ [see (3.7)] $\qquad$ (c) $x^3 - 7x^2 + 14x - 8 = 0$
(b) $x^3 + 2x^2 - 4x - 8 = 0$ $\qquad$ (d) $x^3 - 3x^2 - 4x = 0$

4 Find a cubic function with roots 7, −2, and 5.

5 Find the equilibrium solution for each of the following models:

(a) $Q_d = Q_s$, $Q_d = 3 - P^2$, $Q_s = 6P - 4$

(b) $Q_d = Q_s$, $Q_d = 8 - P^2$, $Q_s = P^2 - 2$

6 The market equilibrium condition, $Q_d = Q_s$, is often expressed in an equivalent alternative form, $Q_d - Q_s = 0$, which has the economic interpretation "excess demand is zero." Does (3.7) represent this latter version of the equilibrium condition? If not, supply an appropriate economic interpretation for (3.7).

3.4 GENERAL MARKET EQUILIBRIUM

The last two sections dealt with models of an isolated market, wherein the Q_d and Q_s of a commodity are functions of the price of that commodity alone. In the actual world, though, no commodity ever enjoys (or suffers) such a hermitic existence; for every commodity, there would normally exist many substitutes and complementary goods. Thus a more realistic depiction of the demand function of a commodity should take into account the effect not only of the price of the commodity itself but also of the prices of most, if not all, of the related commodities. The same also holds true for the supply function. Once the prices of other commodities are brought into the picture, however, the structure of the model itself must be broadened so as to be able to yield the equilibrium values of these other prices as well. As a result, the price and quantity variables of multiple commodities must enter endogenously into the model en masse.

In an isolated-market model, the equilibrium condition consists of only one equation, $Q_d = Q_s$, or $E \equiv Q_d - Q_s = 0$, where E stands for excess demand. When several interdependent commodities are simultaneously considered, equilibrium would require the absence of excess demand for each and every commodity included in the model, for if so much as *one* commodity is faced with an excess demand, the price adjustment of that commodity will necessarily affect the quantities demanded and quantities supplied of the remaining commodities, thereby causing price changes all around. Consequently, the equilibrium condition of an n-commodity market model will involve n equations, one for each commodity, in the form

(3.11) $$E_i \equiv Q_{di} - Q_{si} = 0 \qquad (i = 1, 2, \ldots, n)$$

If a solution exists, there will be a set of prices $\bar{P}_i$ and corresponding quantities $\bar{Q}_i$ such that all the n equations in the equilibrium condition will be simultaneously satisfied.

Two-Commodity Market Model

To illustrate the problem, let us discuss a simple model in which only two commodities are related to each other. For simplicity, the demand and supply functions of both commodities are assumed to be linear. In parametric terms, such a model can be written as

(3.12)
$$\begin{aligned}
&Q_{d1} - Q_{s1} = 0 \\
&Q_{d1} = a_0 + a_1 P_1 + a_2 P_2 \\
&Q_{s1} = b_0 + b_1 P_1 + b_2 P_2 \\
&Q_{d2} - Q_{s2} = 0 \\
&Q_{d2} = \alpha_0 + \alpha_1 P_1 + \alpha_2 P_2 \\
&Q_{s2} = \beta_0 + \beta_1 P_1 + \beta_2 P_2
\end{aligned}$$

where the a and b coefficients pertain to the demand and supply functions of the first commodity, and the α and β coefficients are assigned to those of the second. We have not bothered to specify the signs of the coefficients, but in the course of analysis certain restrictions will emerge as a prerequisite to economically sensible results. Also, in a subsequent numerical example, some comments will be made on the specific signs to be given the coefficients.

As a first step toward the solution of this model, we can again resort to elimination of variables. By substituting the second and third equations into the first (for the first commodity) and the fifth and sixth equations into the fourth (for the second commodity), the model is reduced to two equations in two variables:

$$\begin{aligned} (a_0 - b_0) + (a_1 - b_1)P_1 + (a_2 - b_2)P_2 &= 0 \\ (\alpha_0 - \beta_0) + (\alpha_1 - \beta_1)P_1 + (\alpha_2 - \beta_2)P_2 &= 0 \end{aligned} \tag{3.13}$$

These represent the two-commodity version of (3.11), after the demand and supply functions have been substituted into the two equilibrium-condition equations.

Although this is a simple system of only two equations, as many as 12 parameters are involved, and algebraic manipulations will prove unwieldy unless some sort of shorthand is introduced. Let us therefore define the shorthand symbols

$$\begin{aligned} c_i &\equiv a_i - b_i \\ \gamma_i &\equiv \alpha_i - \beta_i \end{aligned} \qquad (i = 0, 1, 2)$$

Then (3.13) becomes—after transposing the c_0 and γ_0 terms to the right-hand side of the equals sign:

$$\begin{aligned} c_1 P_1 + c_2 P_2 &= -c_0 \\ \gamma_1 P_1 + \gamma_2 P_2 &= -\gamma_0 \end{aligned} \tag{3.13'}$$

which may be solved by further elimination of variables. From the first equation, it can be found that $P_2 = -(c_0 + c_1 P_1)/c_2$. Substituting this into the second equation and solving, we get

$$\bar{P}_1 = \frac{c_2\gamma_0 - c_0\gamma_2}{c_1\gamma_2 - c_2\gamma_1} \tag{3.14}$$

Note that $\bar{P}_1$ is entirely expressed, as a solution value should be, in terms of the data (parameters) of the model. By a similar process, the equilibrium price of the second commodity is found to be

$$\bar{P}_2 = \frac{c_0\gamma_1 - c_1\gamma_0}{c_1\gamma_2 - c_2\gamma_1} \tag{3.15}$$

For these two values to make sense, however, certain restrictions should be imposed on the model. First, since division by zero is undefined, we must require the common denominator of (3.14) and (3.15) to be nonzero, that is, $c_1\gamma_2 \neq c_2\gamma_1$. Second, to assure positivity, the numerator must have the same sign as the denominator.

The equilibrium prices having been found, the equilibrium quantities $\bar{Q}_1$ and $\bar{Q}_2$ can readily be calculated by substituting (3.14) and (3.15) into the second (or third) equation and the fifth (or sixth) equation of (3.12). These solution values will naturally also be expressed in terms of the parameters. (Their actual calculation is left to you as an exercise.)

Numerical Example

Suppose that the demand and supply functions are numerically as follows:

$$
\begin{aligned}
Q_{d1} &= 10 - 2P_1 + P_2 \\
Q_{s1} &= -2 + 3P_1 \\
Q_{d2} &= 15 + P_1 - P_2 \\
Q_{s2} &= -1 \qquad + 2P_2
\end{aligned}
\tag{3.16}
$$

What will be the equilibrium solution?

Before answering the question, let us take a look at the numerical coefficients. For each commodity, Q_{si} is seen to depend on P_i alone, but Q_{di} is shown as a function of both prices. Note that while P_1 has a negative coefficient in Q_{d1}, as we would expect, the coefficient of P_2 is positive. The fact that a rise in P_2 tends to raise Q_{d1} suggests that the two commodities are substitutes for each other. The role of P_1 in the Q_{d2} function has a similar interpretation.

With these coefficients, the shorthand symbols c_i and γ_i will take the following values:

$$
\begin{aligned}
c_0 &= 10 - (-2) = 12 \qquad & c_1 &= -2 - 3 = -5 \qquad & c_2 &= 1 - 0 = 1 \\
\gamma_0 &= 15 - (-1) = 16 \qquad & \gamma_1 &= 1 - 0 = 1 \qquad & \gamma_2 &= -1 - 2 = -3
\end{aligned}
$$

By direct substitution of these into (3.14) and (3.15), we obtain

$$\bar{P}_1 = \tfrac{52}{14} = 3\tfrac{5}{7} \qquad \text{and} \qquad \bar{P}_2 = \tfrac{92}{14} = 6\tfrac{4}{7}$$

And the further substitution of $\bar{P}_1$ and $\bar{P}_2$ into (3.16) will yield

$$\bar{Q}_1 = \tfrac{64}{7} = 9\tfrac{1}{7} \qquad \text{and} \qquad \bar{Q}_2 = \tfrac{85}{7} = 12\tfrac{1}{7}$$

Thus all the equilibrium values turn out positive, as required. In order to preserve the exact values of $\bar{P}_1$ and $\bar{P}_2$ to be used in the further calculation of $\bar{Q}_1$ and $\bar{Q}_2$, it is advisable to express them as fractions rather than decimals.

Could we have obtained the equilibrium prices graphically? The answer is yes. From (3.13), it is clear that a two-commodity model can be summarized by two equations in two variables P_1 and P_2. With known numerical coefficients, both equations can be plotted in the P_1P_2 coordinate plane, and the intersection of the two curves will then pinpoint $\bar{P}_1$ and $\bar{P}_2$.

n-Commodity Case

The above discussion of the multicommodity market has been limited to the case of two commodities, but it should be apparent that we are already moving from *partial-equilibrium* analysis in the direction of *general-equilibrium* analysis. As more commodities enter into a model, there will be more variables and more equations, and the equations will get longer and more complicated. If all the commodities in an economy are included in a comprehensive market model, the result will be a Walrasian type of general-equilibrium model, in which the excess demand for every commodity is considered to be a function of the prices of all the commodities in the economy.

Some of the prices may, of course, carry zero coefficients when they play no role in the determination of the excess demand of a particular commodity; e.g., in the excess-demand function of pianos the price of popcorn may well have a zero coefficient. In general, however, with n commodities in all, we may express the demand and supply functions as follows (using Q_{di} and Q_{si} as function symbols in place of f and g):

$$(3.17) \qquad \begin{aligned} Q_{di} &= Q_{di}(P_1, P_2, \ldots, P_n) \\ Q_{si} &= Q_{si}(P_1, P_2, \ldots, P_n) \end{aligned} \qquad (i = 1, 2, \ldots, n)$$

In view of the index subscript, these two equations represent the totality of the $2n$ functions which the model contains. (These functions are not necessarily linear.) Moreover, the equilibrium condition is itself composed of a set of n equations,

$$(3.18) \qquad Q_{di} - Q_{si} = 0 \qquad (i = 1, 2, \ldots, n)$$

When (3.18) is added to (3.17), the model becomes complete. You should therefore count a total of $3n$ equations.

Upon substitution of (3.17) into (3.18), however, the model can be reduced to a set of n simultaneous equations only:

$$Q_{di}(P_1, P_2, \ldots, P_n) - Q_{si}(P_1, P_2, \ldots, P_n) = 0 \qquad (i = 1, 2, \ldots, n)$$

Besides, inasmuch as $E_i \equiv Q_{di} - Q_{si}$, where E_i is necessarily also a function of all the n prices, the above set of equations may be written alternatively as

$$E_i(P_1, P_2, \ldots, P_n) = 0 \qquad (i = 1, 2, \ldots, n)$$

Solved simultaneously, these n equations will determine the n equilibrium prices $\bar{P}_i$—if a solution does indeed exist. And then the $\bar{Q}_i$ may be derived from the demand or supply functions.

Solution of a General-Equation System

If a model comes equipped with numerical coefficients, as in (3.16), the equilibrium values of the variables will be in numerical terms, too. On a more general level, if a model is expressed in terms of parametric constants, as in (3.12), the equilibrium values will also involve parameters and will hence appear as "formulas," as exemplified by (3.14) and (3.15). If, for greater generality, even the function forms are left unspecified in a model, however, as in (3.17), the manner of expressing the solution values will of necessity be exceedingly general as well.

Drawing upon our experience in parametric models, we know that a solution value is always an expression in terms of the parameters. For a general-function model containing, say, a total of m parameters $(a_1, a_2, \ldots, a_m)$—where m is not necessarily equal to n—the n equilibrium prices can therefore be expected to take the general analytical form of

$$\bar{P}_i = \bar{P}_i(a_1, a_2, \ldots, a_m) \qquad (i = 1, 2, \ldots, n) \tag{3.19}$$

This is a symbolic statement to the effect that the solution value of *each* variable (here, price) is a function of the set of all parameters of the model. As this is a very general statement, it really does not give much detailed information about the solution. But in the general analytical treatment of some types of problem, even this seemingly uninformative way of expressing a solution will prove of use, as will be seen in a later chapter.

Writing such a solution is an easy task. But an important catch exists: the expression in (3.19) can be justified if and only if a *unique* solution does indeed exist, for then and only then can we map the ordered m-tuple $(a_1, a_2, \ldots, a_m)$ into a determinate value for each price $\bar{P}_i$. Yet, unfortunately for us, there is no a priori reason to presume that every model will automatically yield a unique solution. In this connection, it needs to be emphasized that the process of "counting equations and unknowns" does not suffice as a test. Some very simple examples should convince us that an equal number of equations and unknowns (endogenous variables) does not necessarily guarantee the existence of a unique solution.

Consider the three simultaneous-equation systems

$$\begin{aligned} x + y &= 8 \\ x + y &= 9 \end{aligned} \tag{3.20}$$

$$\begin{aligned} 2x + y &= 12 \\ 4x + 2y &= 24 \end{aligned} \tag{3.21}$$

$$\begin{aligned} 2x + 3y &= 58 \\ y &= 18 \\ x + y &= 20 \end{aligned} \tag{3.22}$$

In (3.20), despite the fact that two unknowns are linked together by exactly two equations, there is nevertheless no solution. These two equations happen to be *inconsistent*, for if the sum of x and y is 8, it cannot possibly be 9 at the same time. In (3.21), another case of two equations in two variables, the two equations are *functionally dependent*, which means that one can be derived from (and is implied by) the other. (Here, the second equation is equal to two times the first equation). Consequently, one equation is redundant and may be dropped from the system, leaving in effect only one equation in two unknowns. The solution will then be the equation $y = 12 - 2x$, which yields not a unique ordered pair $(\bar{x}, \bar{y})$ but an infinite number of them, including (0, 12), (1, 10), (2, 8), etc., all of which satisfy that equation. Lastly, the case of (3.22) involves more equations than unknowns, yet the ordered pair (2, 18) does constitute the unique solution to it. The reason is that, in view of the existence of functional dependence among the equations (the first is equal to the second plus twice the third), we have in effect only two independent, consistent equations in two variables.

These simple examples should suffice to convey the importance of *consistency* and *functional independence* as the two prerequisites for application of the process of counting equations and unknowns. In general, in order to apply that process, make sure that (1) the satisfaction of any one equation in the model will not preclude the satisfaction of another and (2) no equation is redundant. In (3.17), for example, the n demand and n supply functions may safely be assumed to be independent of one another, each being derived from a different source—each demand from the decisions of a group of consumers, and each supply from the decisions of a group of firms. Thus each function serves to describe one facet of the market situation, and none is redundant. Mutual consistency may perhaps also be assumed. In addition, the equilibrium-condition equations in (3.18) are also independent and presumably consistent. Therefore the analytical solution as written in (3.19) can in general be considered justifiable.*

For simultaneous-equation models, there exist systematic methods of testing the existence of a unique (or determinate) solution. These would involve, for linear models, an application of the concept of *determinants*, to be introduced in Chap. 5. In the case of nonlinear models, such a test would also require a knowledge of so-called "partial derivatives" and a special type of determinant called the *Jacobian determinant*, which will be discussed in Chaps. 7 and 8.

EXERCISE 3.4

1 Work out the step-by-step solution of (3.13′), thereby verifying the results in (3.14) and (3.15).

2 Rewrite (3.14) and (3.15) in terms of the original parameters of the model in (3.12).

3 The demand and supply functions of a two-commodity market model are as follows:

$$Q_{d1} = 18 - 3P_1 + P_2 \qquad Q_{d2} = 12 + P_1 - 2P_2$$
$$Q_{s1} = -2 + 4P_1 \qquad Q_{s2} = -2 + 3P_2$$

Find $\bar{P}_i$ and $\bar{Q}_i$ $(i = 1, 2)$. (Use fractions rather than decimals.)

* This is essentially the way that Léon Walras approached the problem of the existence of a general market equilibrium. In the modern literature, there can be found a number of sophisticated mathematical proofs of the existence of a competitive market equilibrium under certain postulated economic conditions. But the mathematics used is advanced. The easiest one to understand is perhaps the proof given in Robert Dorfman, Paul A. Samuelson, and Robert M. Solow, *Linear Programming and Economic Analysis*, McGraw-Hill Book Company, New York, 1958, chapter 13, which you should read *after* having studied Part 6 of the present volume.

3.5 EQUILIBRIUM IN NATIONAL-INCOME ANALYSIS

Even though the discussion of static analysis has hitherto been restricted to *market models* in various guises—linear and nonlinear, one-commodity and multicommodity, specific and general—it, of course, has applications in other areas of economics also. As a simple example, we may cite the familiar Keynesian national-income model,

$$\begin{aligned} Y &= C + I_0 + G_0 \\ C &= a + bY \end{aligned} \qquad (a > 0, \quad 0 < b < 1) \tag{3.23}$$

where Y and C stand for the endogenous variables national income and consumption expenditure, respectively, and I_0 and G_0 represent the exogenously determined investment and government expenditures. The first equation is an equilibrium condition (national income = total expenditure). The second, the consumption function, is behavioral. The two parameters in the consumption function, a and b, stand for the autonomous consumption expenditure and the marginal propensity to consume, respectively.

It is quite clear that these two equations in two endogenous variables are neither functionally dependent upon, nor inconsistent with, each other. Thus we would be able to find the equilibrium values of income and consumption expenditure, $\bar{Y}$ and $\bar{C}$, in terms of the parameters a and b and the exogenous variables I_0 and G_0.

Substitution of the second equation into the first will reduce (3.23) to a single equation in one variable, Y:

$$Y = a + bY + I_0 + G_0$$

or $$(1 - b)Y = a + I_0 + G_0$$

Thus the solution value of Y (equilibrium national income) is

$$\bar{Y} = \frac{a + I_0 + G_0}{1 - b} \tag{3.24}$$

which, it should be noted, is expressed entirely in terms of the parameters and exogenous variables, the given data of the model. Putting (3.24) into the second equation of (3.23) will then yield the equilibrium level of consumption expenditure:

$$\begin{aligned} \bar{C} = a + b\bar{Y} &= a + \frac{b(a + I_0 + G_0)}{1 - b} \\ &= \frac{a(1 - b) + b(a + I_0 + G_0)}{1 - b} = \frac{a + b(I_0 + G_0)}{1 - b} \end{aligned} \tag{3.25}$$

which is again expressed entirely in terms of the given data.

Both $\bar{Y}$ and $\bar{C}$ have the expression $(1 - b)$ in the denominator; thus a restriction $b \neq 1$ is necessary, to avoid division by zero. Since b, the marginal propensity to consume, has been assumed to be a positive fraction, this restriction is automatically satisfied. For $\bar{Y}$ and $\bar{C}$ to be positive, moreover, the numerators in

(3.24) and (3.25) must be positive. Since the exogenous expenditures I_0 and G_0 are normally positive, as is the parameter a (the vertical intercept of the consumption function), the sign of the numerator expressions will work out, too.

As a check on our calculation, we can add the $\bar{C}$ expression in (3.25) to $(I_0 + G_0)$ and see whether the sum is equal to the $\bar{Y}$ expression in (3.24). If so, the $\bar{C}$ and $\bar{Y}$ values do satisfy the equilibrium condition, and the solution is valid.

This model is obviously one of extreme simplicity and crudity, but other models of national-income determination, in varying degrees of complexity and sophistication, can be constructed as well. In each case, however, the principles involved in the construction and analysis of the model are identical with those already discussed. For this reason, we shall not go into further illustrations here. A more comprehensive national-income model, involving the simultaneous equilibrium of the money market and the goods market, will be discussed in Sec. 8.6 below.

EXERCISE 3.5

1 Given the following model:

$Y = C + I_0 + G_0$

$C = a + b(Y - T) \qquad (a > 0, \quad 0 < b < 1) \qquad [T\text{: taxes}]$

$T = d + tY \qquad (d > 0, \quad 0 < t < 1) \qquad [t\text{: income tax rate}]$

(a) How many endogenous variables are there?

(b) Find $\bar{Y}$, $\bar{T}$, and $\bar{C}$.

2 Let the national-income model be:

$Y = C + I_0 + G$

$C = a + b(Y - T_0) \qquad (a > 0, \quad 0 < b < 1)$

$G = gY \qquad (0 < g < 1)$

(a) Identify the endogenous variables.

(b) Give the economic meaning of the parameter g.

(c) Find the equilibrium national income.

(d) What restriction on the parameters is needed for a solution to exist?

3 Find $\bar{Y}$ and $\bar{C}$ from the following:

$Y = C + I_0 + G_0$

$C = 25 + 6Y^{1/2}$

$I_0 = 16$

$G_0 = 14$