chapter 1.

Introduction

1.1 DYNAMIC PHENOMENA

The term *dynamic* refers to phenomena that produce time-changing patterns, the characteristics of the pattern at one time being interrelated with those at other times. The term is nearly synonymous with *time-evolution* or *pattern of change.* It refers to the unfolding of events in a continuing evolutionary process.

Nearly all observed phenomena in our daily lives or in scientific investigation have important dynamic aspects. Specific examples may arise in (a) a physical system, such as a traveling space vehicle, a home heating system, or in the mining of a mineral deposit; (b) a social system, such as the movement within an organizational hierarchy, the evolution of a tribal class system, or the behavior of an economic structure; or (c) a life system, such as that of genetic transference, ecological decay, or population growth. But while these examples illustrate the pervasiveness of dynamic situations and indicate the potential value of developing the facility for representing and analyzing dynamic behavior, it must be emphasized that the general concept of dynamics transcends the particular origin or setting of the process.

Many dynamic systems can be understood and analyzed intuitively, without resort to mathematics and without development of a general theory of dynamics. Indeed, we often deal quite effectively with many simple dynamic situations in our daily lives. However, in order to approach unfamiliar complex situations efficiently, it is necessary to proceed systematically. Mathematics can provide the required economy of language and conceptual framework.

With this view, the term *dynamics* soon takes on somewhat of a dual meaning. It is, first, as stated earlier, a term for the time-evolutionary phenomena in the world about us, and, second, it is a term for that part of mathematical science that is used for the representation and analysis of such phenomena. In the most profound sense the term refers simultaneously to both aspects: the real, the abstract, and the interplay between them.

Although there are endless examples of interesting dynamic situations arising in a spectrum of areas, the number of corresponding general forms for mathematical representation is relatively small. Most commonly, dynamic systems are represented mathematically in terms of either differential or difference equations. Indeed, this is so much the case that, in terms of pure mathematical content, at least the elementary study of dynamics is almost synonymous with the theory of differential and difference equations. It is these equations that provide the structure for representing time linkages among variables.

The use of either differential or difference equations to represent dynamic behavior corresponds, respectively, to whether the behavior is viewed as occurring in continuous or discrete time. Continuous time corresponds to our usual conception, where time is regarded as a continuous variable and is often viewed as flowing smoothly past us. In mathematical terms, continuous time of this sort is quantified in terms of the continuum of real numbers. An arbitrary value of this continuous time is usually denoted by the letter t. Dynamic behavior viewed in continuous time is usually described by differential equations, which relate the derivatives of a dynamic variable to its current value.

Discrete time consists of an ordered sequence of points rather than a continuum. In terms of applications, it is convenient to introduce this kind of time when events and consequences either occur or are accounted for only at discrete time periods, such as daily, monthly, or yearly. When developing a population model, for example, it may be convenient to work with yearly population changes rather than continuous time changes. Discrete time is usually labeled by simply indexing, in order, the discrete time points, starting at a convenient reference point. Thus time corresponds to integers 0, 1, 2, and so forth, and an arbitrary time point is usually denoted by the letter k. Accordingly, dynamic behavior viewed in discrete time is usually described by equations relating the value of a variable at one time to the values at adjacent times. Such equations are called *difference equations.*

1.2 MULTIVARIABLE SYSTEMS

The term *system,* as applied to general analysis, was originated as a recognition that meaningful investigation of a particular phenomenon can often only be achieved by explicitly accounting for its environment. The particular variables of interest are likely to represent simply one component of a complex, consisting of perhaps several other components. Meaningful analysis must consider the entire system and the relations among its components. Accordingly, mathematical models of systems are likely to involve a large number of interrelated variables—and this is emphasized by describing such situations as *multivariable systems*. Some examples illustrating the pervasiveness and importance of multivariable phenomena arise in consideration of (a) the migration patterns of population between various geographical areas, (b) the simultaneous interaction of various individuals in an economic system, or (c) the various age groups in a growing population.

The ability to deal effectively with large numbers of interrelated variables is one of the most important characteristics of mathematical system analysis. It is necessary therefore to develop facility with techniques that help one clearly think about and systematically manipulate large numbers of simultaneous relations. For one's own thinking purposes, in order to understand the essential elements of the situation, one must learn, first, to view the whole set of relations as a unit, suppressing the details; and, second, to see the important detailed interrelations when required. For purposes of manipulation, with the primary objective of computation rather than furthering insight, one requires a systematic and efficient representation.

There are two main methods for representing sets of interrelations. The first is vector notation, which provides an efficient representation both for computation and for theoretical development. By its very nature, vector notation suppresses detail but allows for its retrieval when required. It is therefore a convenient, effective, and practical language. Moreover, once a situation is cast in this form, the entire array of theoretical results from linear algebra is available for application. Thus, this language is also well matched to mathematical theory.

The second technique for representing interrelations between variables is by use of diagrams. In this approach the various components of a system are represented by points or blocks, with connecting lines representing relations between the corresponding components. This representation is exceedingly helpful for visualization of essential structure in many complex situations; however, it lacks the full analytical power of the vector method. It is for this reason that, although both methods are developed in this book, primary emphasis is placed on the vector approach.

Most situations that we investigate are both dynamic and multivariable. They are, accordingly, characterized by several variables, each changing with time and each linked through time to other variables. Indeed, this combination of multivariable and time-evolutionary structure characterizes the setting of the modern theory of dynamic systems.

That most dynamic systems are both time-evolutionary and multivariable implies something about the nature of the mathematics that forms the basis for their analysis. The mathematical tools are essentially a combination of differential (or difference) equations and vector algebra. The differential (or difference) equations provide the element of dynamics, and the vector algebra provides the notation for multivariable representation. The combination and interplay between these two branches of mathematics provides the basic foundation for all analysis in this book. It is for this reason that this introductory chapter is followed first by a chapter on differential and difference equations and then by a chapter on matrix algebra.

1.3 A CATALOG OF EXAMPLES

As in all areas of problem formulation and analysis, the process of passing from a "real world" dynamic situation to a suitable abstraction in terms of a mathematical model requires an expertise that is refined only through experience. In any given application there is generally no single "correct" model; rather, the degree of detail, the emphasis, and the choice of model form are subject to the discretionary choice of the analyst. There are, however, a number of models that are considered "classic" in that they are well-known and generally accepted. These classic models serve an important role, not only as models of the situation that they were originally intended to represent, but also as examples of the degree of clarity and reality one should strive to achieve in new situations. A proficient analyst usually possesses a large mental catalog of these classic models that serve as valuable reference points—as well-founded points of departure.

The examples in this section are in this sense all classic, and as such can form the beginnings of a catalog for the reader. The catalog expands as one works his way through succeeding chapters, and this growth of well-founded examples with known properties should be one of the most important objectives of one's study. A diverse catalog enriches the process of model development.

The first four examples are formulated in discrete time and are, accordingly, defined by difference equations. The last two are defined in continuous time and thus result in differential equations. It will be apparent from a study of the examples that the choice to develop a continuous-time or a discrete-time model of a specific phenomenon is somewhat arbitrary. The choice is usually resolved on the basis of data availability, analytical tractability, established convention in the application area, or simply personal preference.

Example 1 (Geometric Growth). A simple growth law, useful in a wide assortment of situations (such as describing the increase in human or other

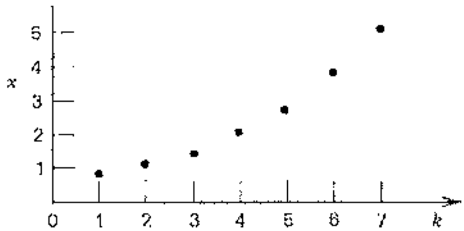

Figure 1.1. Geometric growth.

populations, the growth of vegetation, accumulated publications in a scientific field, consumption of raw materials, the accumulation of interest on a loan, etc.), is the linear law described by the difference equation

$$x(k+1) = ax(k)$$

The value $x(k)$ represents the magnitude of the variable (e.g., population) at time instant k. The parameter a is a constant that determines the rate of growth. For positive growth, the value of a must be greater than unity—then each successive magnitude is a fixed factor larger than its predecessor.

If an initial magnitude is given, say $x(0)=1$, the successive values can be found recursively. In particular, it is easy to see that $x(1)=a$, $x(2)=a^2$, and, in general, $x(k)=a^k$ for $k=0, 1, 2, \ldots$. A typical pattern of growth resulting from this model is shown in Fig. 1.1.

The growth pattern resulting from this simple linear model is referred to as *geometric growth* since the values grow as the terms of a geometric series. This form of growth pattern has been found to agree closely with empirical data in many situations, and there is often strong accompanying theoretical justification for the model, at least over a range of values.

Example 2 (Cohort Population Model). For many purposes (particularly in populations where the level of reproductive activity is nonuniform over a normal lifetime) the simple growth model given above is inadequate for comprehensive analysis of population change. More satisfactory models take account of the age distribution within the population. The classical model of this type is referred to as a *cohort population model.*

The population is divided into age groups (or cohorts) of equal age span, say five years. That is, the first group consists of all those members of the population between the ages of zero and five years, the second consists of those between five and ten years, and so forth. The cohort model itself is a discrete-time dynamic system with the duration of a single time period corresponding to the basic cohort span (five years in our example). By assuming that the male and female populations are identical in distribution, it is possible to simplify the model by considering only the female population. Let $x_i(k)$ be the (female) population of the ith age group at time period k. The groups are indexed sequentially from 0 through n, with 0 representing the lowest age group and n the largest. To describe system behavior, it is only necessary to describe how these numbers change during one time period.

First, aside from the possibility of death, which will be considered in a moment, it is clear that during one time period the cohorts in the ith age group simply move up to the $(i+1)$th age group. To account for the death rate of individuals within a given age group, this upward progression is attenuated by a survival factor. The net progression can be described by the simple equations

$$x_{i+1}(k+1)=\beta_i x_i(k), \qquad i=0,1,\ldots,n-1 \tag{1-1}$$

where β_i is the survival rate of the ith age group during one period. The factors β_i can be determined statistically from actuarial tables.

The only age group not determined by the equation above is $x_0(k+1)$, the group of individuals born during the last time period. They are offspring of the population that existed in the previous time period. The number in this group depends on the birth rate of each of the other cohort groups, and on how large each of these groups was during the previous period. Specifically,

$$x_0(k+1)=\alpha_0 x_0(k)+\alpha_1 x_1(k)+\alpha_2 x_2(k)+\cdots+\alpha_n x_n(k) \tag{1-2}$$

where α_i is the birth rate of the ith age group (expressed in number of female offspring per time period per member of age group i). The factor α_i also can be usually determined from statistical records.

Together Eqs. (1-1) and (1-2) define the system equations, determining how $x_i(k+1)$'s are found from $x_i(k)$'s. This is an excellent example of the combination of dynamics and multivariable system structure. The population system is most naturally visualized in terms of the variables representing the population levels of the various cohort groups, and thus it is a multivariable system. These variables are linked dynamically by simple difference equations, and thus the whole can be regarded as a composite of difference equations and multivariable structure.

Example 3 (National Economics). There are several simple models of national economic dynamics.* We present one formulated in discrete time, where the time between periods is usually taken as quarters of full years. At each time period there are four variables that define the model. They are

$$Y(k)=\text{National Income or National Product}$$
$$C(k)=\text{Consumption}$$
$$I(k)=\text{Investment}$$
$$G(k)=\text{Government Expenditure}$$

* See the notes and references for Sect. 4.8, at the end of Chapter 4.

The variable Y is defined to be the National Income: the total amount earned during a period by all individuals in the economy. Alternatively, but equivalently, Y can be defined as the National Product: the total value of goods and services produced in the economy during the period. Consumption C is the total amount spent by individuals for goods and services. It is the total of every individual's expenditure. The Investment I is the total amount invested in the period. Finally, G is the total amount spent by government during the period, which is equal to the government's current revenue. The basic national accounting equation is

$$Y(k) = C(k) + I(k) + G(k) \tag{1-3}$$

From an income viewpoint, the equation states that total individual income must be divided among consumption of goods and services, investment, or payments to the government. Alternatively, from a national product viewpoint, the total aggregate of goods and services produced must be divided among individual consumption, investment, or government consumption.

In addition to this basic definitional equation, two relationships are introduced that represent assumptions on the behavior of the economy. First, it is assumed that consumption is a fixed fraction of national income. Thus,

$$C(k) = mY(k) \tag{1-4}$$

for some m. The number m, which is restricted to the values $0 < m < 1$, is referred to as the *marginal propensity to consume.* This equation assumes that on the average individuals tend to consume a fixed portion of their income.

The second assumption concerning how the economy behaves relates to the influence of investment. The general effect of investment is to increase the productive capacity of the nation. Thus, present investment will increase national income (or national product) in future years. Specifically, it is assumed that the increase in national income is proportional to the level of investment. Or,

$$Y(k+1) - Y(k) = rI(k) \tag{1-5}$$

The constant r is the *growth factor,* and it is assumed that $r > 0$.

The set of equations (1-3), (1-4), and (1-5) defines the operation of the economy. Of the three equations, only the last is dynamic. The first two, (1-3) and (1-4), are *static,* expressing relationships among the variables that hold at every k. These two static equations can be used to eliminate two variables from the model. Starting with

$$Y(k) = C(k) + I(k) + G(k)$$

substitution of (1-4) produces

$$Y(k) = mY(k) + I(k) + G(k)$$

Substitution of (1-5) then produces

$$Y(k)=mY(k)+\frac{Y(k+1)-Y(k)}{r}+G(k)$$

Rearrangement leads to the final result:

$$Y(k+1)=[1+r(1-m)]Y(k)-rG(k) \tag{1-6}$$

The quantity $G(k)$ appears as an input to the system. If $G(k)$ were held equal to zero, the model would be identical to the first-order (geometric) growth model discussed earlier.

Example 4 (Exponential Growth). The continuous-time version of the simple first-order growth model (the analog of geometric growth) is defined by the differential equation

$$\frac{dx(t)}{dt}=rx(t)$$

The growth parameter r can be any real value, but for (increasing) growth it must be greater than zero. The solution to the equation is found by writing it in the form

$$\frac{1}{x(t)}\frac{dx(t)}{dt}=r$$

Both sides can then be integrated with respect to t to produce

$$\log x(t)=rt+\log c=\log e^{rt}+\log c$$

where c is an arbitrary constant. Taking the antilog yields

$$x(t)=ce^{rt}$$

Finally, by setting $t=0$, it is seen that $x(0)=c$, so the solution can be written

$$x(t)=x(0)e^{rt}$$

This is the equation of *exponential growth*. The solution is sketched for various values of r in Fig. 1.2.

The pattern of solutions is similar to that of geometric growth shown in Fig. 1.1 in Sect. 1.6. Indeed, a series of values from the continuous-time solution at equally spaced time points make up a geometric growth pattern.

Example 5 (Newton's Laws). A wealth of dynamic system examples is found in mechanical systems governed by Newton's laws. In fact, many of the general techniques for dynamic system analysis were originally motivated by such applications. As a simple example, consider motion in a single dimension—of, say, a street car or cable car of mass M moving along a straight track. Suppose

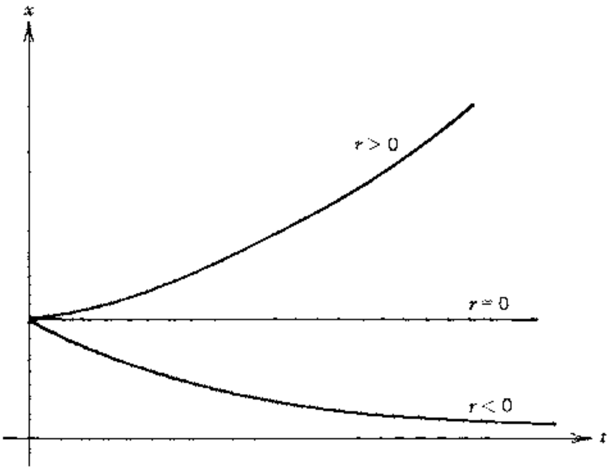

Figure 1.2. Exponential growth.

the position of the car along the track at time t is denoted by $y(t)$, and the force applied to the street car, parallel to the track, is denoted by $u(t)$. Newton's second law says that force is equal to mass times acceleration, or, mathematically,

$$u(t) = M\frac{d^2y}{dt^2}$$

Therefore, the motion is defined by a second-order differential equation.

A more detailed model would, of course, have many other variables and equations to account for spring action, rocking, and bouncing motion, and to account for the fact that forces are applied only indirectly to the main bulk through torque on the wheels or from a friction grip on a cable. The degree of detail constructed into the model would depend on the use to which the model were to be put.

Figure 1.3. Cable car.

Example 6 (Goats and Wolves). Imagine an island populated primarily by goats and wolves. The goats survive by eating the island vegetation. The wolves survive by eating the goats.

The modeling of this kind of population system, referred to as a predator–prey system, goes back to Volterra in response to the observation that populations of species often oscillated. In our example, goats would first be plentiful but wolves rare, and then wolves would be plentiful but goats rare. Volterra described the situation in the following way.

Let

$$N_1(t) = \text{number of goats at time } t$$
$$N_2(t) = \text{number of wolves at time } t$$

The proposed model is then

$$\frac{dN_1(t)}{dt} = aN_1(t) - bN_1(t)N_2(t)$$

$$\frac{dN_2(t)}{dt} = -cN_2(t) + dN_1(t)N_2(t)$$

where the constants a, b, c, and d are all positive.

This model, which is the archetype of predator–prey models, has a simple biological interpretation. In the absence of wolves [$N_2(t) = 0$], the goat population is governed by simple exponential growth, with growth factor a. The goats thrive on the island vegetation. In the absence of goats [$N_1(t) = 0$], on the other hand, the wolf population is governed by exponential decline, declining at a rate $-c$. This interpretation accounts for the first terms on the right-hand side of the differential equations.

When both goats and wolves are present on the island, there are encounters between the two groups. Under an assumption of random movement, the frequency of encounters is proportional to the product of the numbers in the two populations. Each encounter decreases the goat population and increases the wolf population. The effect of these encounters is accounted for by the second terms in the differential equations.

1.4 THE STAGES OF DYNAMIC SYSTEM ANALYSIS

The principal objectives of an analysis of a dynamic system are as varied as the range of possible application areas. Nevertheless, it is helpful to distinguish four (often overlapping) stages of dynamic analysis: representation of phenomena, generation of solutions, exploration of structural relations, and control or modification. Most analyses emphasize one or two of these stages, with the others having been completed previously or lying beyond the reach of current technique.

A recognition of these four stages helps motivate the assortment of theoretical principles associated with the mathematics of dynamic systems, for there is, naturally, great interplay between general theory and the analysis of given situations. On the one hand, the objectives for an analysis are strongly influenced by available theory, and, on the other hand, development of new theory is often motivated by the desire to conduct deeper analyses.

Representation

One of the primary objectives of the use of mathematics in complex dynamic systems is to obtain a mathematical representation of the system, and this is the first stage of analysis. The process of obtaining the representation is often referred to as *modeling*, and the final product a *model*. This stage is closely related to the sciences, for the development of a suitable model amounts to the employment or development of scientific theory. The theory employed in any given model may be well-founded and generally accepted, or it may be based only on one analyst's hypothesized relationships. A complex model will often have both strong and weak components. But in any case the model description is an encapsulation of a scientific theory.

Development of a meaningful representation of a complex system requires more than just scientific knowledge. The end product is likely to be most meaningful if one understands the theory of dynamic systems as well as the relevant scientific theory. Only then is it possible to assess, at least in qualitative terms, the dynamic significance of various assumptions, and thereby build a model that behaves in a manner consistent with intuitive expectations.

Generation of Solutions

The most direct use of a dynamic model is the generation of a specific solution to its describing equations. The resulting time pattern of the variables then can be studied for various purposes.

A specific solution can sometimes be found in analytical form, but more often it is necessary to generate specific solutions numerically by use of a calculator or digital computer—a process commonly referred to as *simulation*. As an example of this direct use of a model, a large cohort model of a nation's population growth can be solved numerically to generate predictions of future population levels, catalogued by age group, sex, and race. The results of such a simulation might be useful for various planning problems. Likewise, a model of the national economy can forecast future economic trends, thereby possibly suggesting the appropriateness of various corrective policies. Or, in the context of any situation, simulation might be used to test the reasonableness of a new model by verifying that a particular solution has the properties usually associated with the underlying phenomena.

It is of course rare that a single solution of a model is adequate for a meaningful analysis. Every model really represents a collection of solutions, each determined by different controlled inputs, different parameter values, and different starting conditions. In the population system, for example, the specific future population level is dependent on national immigration policy, on the birth rates in future years, and on the assumed level of current population. One may therefore find that it is necessary to generate solutions corresponding to various combinations of assumptions in order to conduct a meaningful analysis of probable future population.

As a general rule, the number of required solutions grows quickly with the number of different parameters and inputs that must be varied independently. Thus, although direct simulation is a flexible concept applicable to quite large and complex systems where analysis is difficult, it is somewhat limited in its capability to explore all ranges of input and parameter values.

Exploration of Structural Relations

Much of the theory of dynamic systems is motivated by a desire to go beyond the stage of simply computing particular solutions of a model to the point of establishing various structural relations as, say, between a certain parameter and its influence on the solution. Such relations are often obtained indirectly through the use of auxiliary concepts of analysis.

The payoff of this type of structural exploration manifests itself in two important and complementary ways. First, it develops intuitive insight into system behavior. With this insight, one is often able to determine the rough outlines of the solution to a complex system almost by inspection, and, more importantly, to foresee the nature of the effects of possible system modifications. But it is important to stress that the value of this insight goes well beyond the mere approximation of a solution. Insight into system behavior is reflected back, as an essential part of the creative process, to refinement of the formulation of the original model. A model will be finally accepted only when one is assured of its reasonableness—both in terms of its structure and in terms of the behavior patterns it generates.

The second payoff of structural exploration is that it often enables one to explicitly calculate relations that otherwise could be deduced only after examination of numerous particular solutions. For example, as is shown in Chapter 5, the natural rate of growth of a cohort population model can be determined directly from its various birth rate and survival rate coefficients, without generating even a single specific growth pattern. This leads, for example, to a specific relationship between changes in birth rates and changes in composite population growth. In a similar fashion, the stability of a complex economic process of price adjustment can often be inferred from its structural form, without generating solutions.

Most of the theoretical development in this book is aimed at revealing relationships of this kind between structure and behavior. By learning this theory we become more than just equation writers and equation solvers. Our analysis is not limited in its application to a particular problem with particular numerical constants, but instead is applicable to whole classes of models; and results from one situation can be readily transferred to another.

Control or Modification

Although study of a particular dynamic situation is sometimes motivated by the simple philosophic desire to understand the world and its phenomena, many analyses have the explicit motivation of devising effective means for changing a system so that its behavior pattern is in some way improved. The means for affecting behavior can be described as being either system modification or control. Modification refers to a change in the system, and hence in its describing equation. This might be a change in various parameter values or the introduction of new interconnective mechanisms. Examples of modification are: a change in the birth rates of a population system, a change of marriage rules in a class society, a change of forecasting procedure in an economic system, a change of promotion rate in an organizational hierarchy, and so forth. Control, on the other hand, generally implies a continuing activity executed throughout the operation of the system. The Federal Reserve Board controls the generation of new money in the economy on a continuing basis, a farmer controls the development of his herd of cattle by controlling the amount of grain they are fed, a pilot controls the behavior of his aircraft continuously, and so forth.

Determination of a suitable modification or control strategy for a system represents the fourth stage of analysis, and generally marks the conclusion of a complete analysis cycle. However, at the completion of the best analyses, the main outlines of the solution should be fairly intuitive—during the course of analysis the intuition should be heightened to a level sufficient to accept the conclusions. Mathematics serves as a language for organized thought, and thought development, not as a machine for generating complexity. The mathematics of dynamic systems is developed to expedite our requests for detail when required, and to enhance our insight into the behavior of dynamic phenomena we encounter in the world.