1

What Is Mathematical Modeling?

We begin this book with a dictionary definition of the word model :

model (n): a miniature representation of something; a pattern of something to be made; an example for imitation or emulation; a description or analogy used to help visualize something (e.g., an atom) that cannot be directly observed; a system of postulates, data and inferences presented as a mathematical description of an entity or state of affairs

This definition suggests that modeling is an activity, a cognitive activity in which we think about and make models to describe how devices or objects of interest behave.
There are many ways in which devices and behaviors can be described. We can use words, drawings or sketches, physical models, computer programs, or mathematical formulas. In other words, the modeling activity can be done in several languages, often simultaneously. Since we are particularly interested in using the language of mathematics to make models, we will refine the definition just given:

mathematical model (n): a representation in mathematical terms of the behavior of real devices and objects

We want to know how to make or generate mathematical representations ormodels, how to validate them, how to use them, and how and when their use is limited. But before delving into these important issues, it is worth talking about why we do mathematical modeling.

1.1 Why Do We Do Mathematical Modeling?

Since the modeling of devices and phenomena is essential to both engineering and science, engineers and scientists have very practical reasons for doing mathematical modeling. In addition, engineers, scientists, and mathematicianswant to experience the sheer joy of formulating and solving mathematical problems.

1.1.1 Mathematical Modeling and the Scientific Method

In an elementary picture of the scientificmethod (see Figure 1.1), we identify a "real world" and a "conceptual world." The external world is the one we call real; here we observe various phenomena and behaviors, whether natural in origin or produced by artifacts. The conceptual world is the world of the mind—where we live when we try to understand what is going on in that real, external world. The conceptual world can be viewed as having three stages: observation, modeling, and prediction.

In the observation part of the scientific method we measure what is happening in the real world. Here we gather empirical evidence and "facts on the ground." Observations may be direct, as when we use our senses, or indirect, in which case some measurements are taken to indicate through some other reading that an event has taken place. For example, we often know a chemical reaction has taken place only by measuring the product of that reaction.

In this elementary view of how science is done, the modeling part is concerned with analyzing the above observations for one of (at least) three reasons. These rationales are about developing: models that describe the behavior or results observed; models that explain why that behavior and results occurred as they did; or models that allow us to predict future behaviors or results that are as yet unseen or unmeasured.

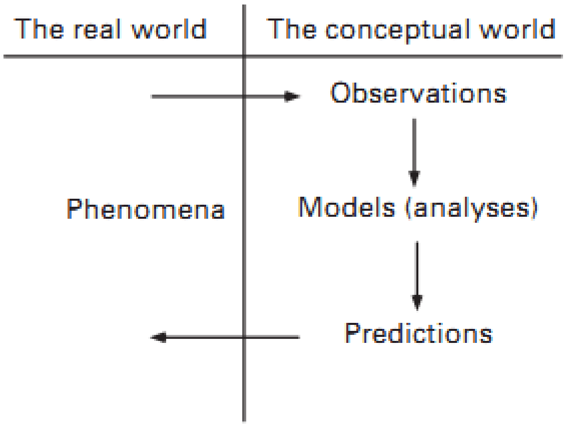

Figure 1.1 An elementary depiction of the *scientific method* that shows how our conceptual models of the world are related to observations made within that real world (Dym and Ivey, 1980).

In the prediction part of the scientific method we exercise our models to tell us what will happen in a yet-to-be-conducted experiment or in an anticipated set of events in the real world. These predictions are then followed by observations that serve either to validate themodel or to suggest reasons that the model is inadequate.

The last point clearly points to the looping, iterative structure apparent in Figure 1.1. It also suggests that modeling is central to all of the conceptual phases in the elementary model of the scientific method.We build models and use them to predict events that can confirm or deny the models. In addition, we can also improve our gathering of empirical data when we use a model to obtain guidance about where to look.

1.1.2 Mathematical Modeling and the Practice of Engineering

Engineers are interested in designing devices and processes and systems. That is, beyond observing how the world works, engineers are interested in creating artifacts that have not yet come to life. As noted by Herbert A. Simon (in The Sciences of the Artificial), “Design is the distinguishing activity of engineering.” Thus, engineers must be able to describe and analyze objects and devices into order to predict their behavior to see if that behavior is what the engineers want. In short, engineers need to model devices and processes if they are going to design those devices and processes.

While the scientific method and engineering design have much in common, there are differences in motivation and approach that are worth mentioning. In the practices of science and of engineering design, models are often applied to predict what will happen in a future situation. In engineering design, however, the predictions are used in ways that have far different consequences than simply anticipating the outcome of an experiment. Every new building or airplane, for example, represents a model-based prediction that the building will stand or the airplane will fly without dire, unanticipated consequences. Thus, beyond simply validating a model, prediction in engineering design assumes that resources of time, imagination, and money can be invested with confidence because the predicted outcome will be a good one.

1.2 Principles of Mathematical Modeling

Mathematical modeling is a principled activity that has both principles behind it and methods that can be successfully applied. The principles are over-arching or meta-principles phrased as questions about the intentions and purposes ofmathematicalmodeling. Thesemeta-principles are almost philosophical in nature.We will now outline the principles, and in the next section we will briefly review some of the methods.

A visual portrayal of the basic philosophical approach is shown in Figure 1.2. These methodological modeling principles are also captured in the following list of questions and answers:

• **Why?** What are we looking for? Identify the need for the model.

• **Find?** What do we want to know? List the data we are seeking.

• **Given?** What do we know? Identify the available relevant data.

• **Assume?** What can we assume? Identify the circumstances that apply.

• **How?** How should we look at this model? Identify the governing physical principles.

• **Predict?** What will ourmodel predict? Identify the equations that will be used, the calculations that will be made, and the answers that will result.

• **Valid?** Are the predictions valid? Identify tests that can be made to validate the model, i.e., is it consistent with its principles and assumptions?

• **Verified?** Are the predictions good? Identify tests that can be made to verify the model, i.e., is it useful in terms of the initial reason it was done?

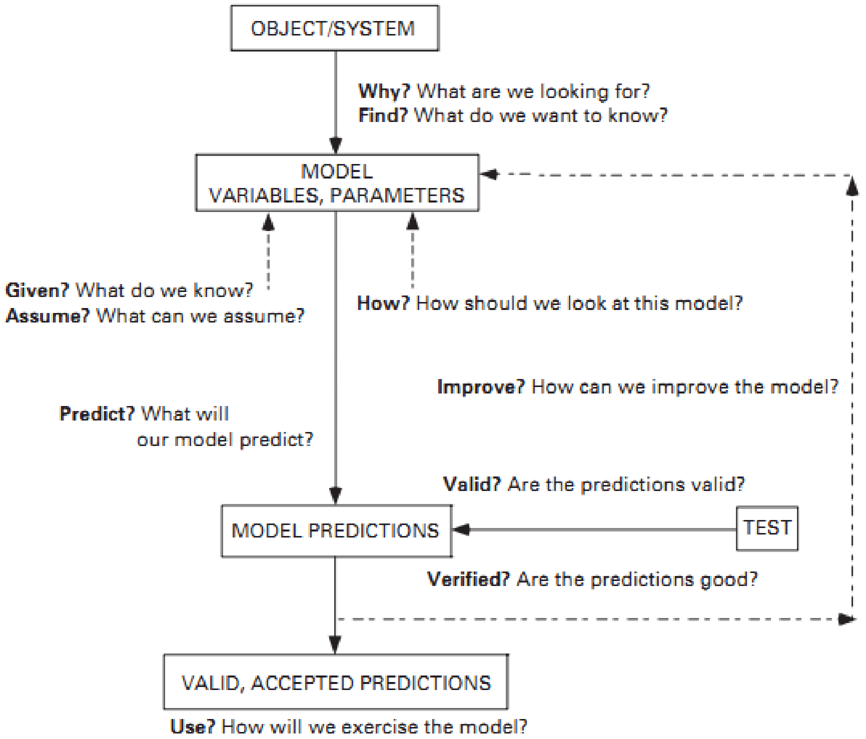

Figure 1.2 A first-order view of *mathematical modeling* that shows how the questions asked in a principled approach to building a model relate to the development of that model (inspired by Carson and Cobelli, 2001).

• **Improve?** Can we improve the model? Identify parameter values that are not adequately known, variables that should have been included, and/or assumptions/restrictions that could be lifted. Implement the iterative loop thatwe can call "model-validate-verify-improve-predict."

• **Use?** Howwillwe exercise themodel?Whatwillwe dowith themodel?

This list of questions and instructions is *not* an algorithm for building a good mathematical model. However, the underlying ideas are key to mathematical modeling, as they are key to problem formulation generally. Thus, we should expect the individual questions to recur often during the modeling process, andwe should regard this list as a fairly general approach to *ways of thinking* about mathematical modeling.

Having a clear picture of why the model is wanted or needed is of prime importance to themodel-building enterprise. Suppose we want to estimate how much power could be generated by a dam on a large river, say a dam located at The Three Gorges on the Yangtze River in Hubei Province in the People's Republic of China. For a first estimate of the available power, we wouldn't need to model the dam's thickness or the strength of its foundation. Its height, on the other hand, would be an essential parameter of a power model, as would some model and estimates of river flow quantities. If, on the other hand, we want to design the actual dam, we would need a model that incorporates all of the dam's physical characteristics (e.g., dimensions, materials, foundations) and relates them to the dam site and the river flow conditions. Thus, defining the task is the first essential step in model formulation.

We then should list what we know—for example, river flow quantities and desired power levels—as a basis for listing the variables or parameters that are as yet unknown. We should also list any relevant assumptions. For example, levels of desired power may be linked to demographic or economic data, so any assumptions made about population and economic growth should be spelled out. Assumptions about the consistency of river flows and the statistics of flooding should also be spelled out.

Which physical principles apply to this model? The mass of the river's water must be conserved, as must its momentum, as the river flows, and energy is both dissipated and redirected as water is allowed to flow through turbines in the dam(and hopefully not spill over the top!). And mass must be conserved, within some undefined system boundary, because dams do accumulate water mass from flowing rivers. There are well-known equations that correspond to these physical principles. They could be used to develop an estimate of dam height as a function of power desired. We can validate the model by ensuring that our equations and calculated results have the proper dimensions, and we can exercise the model against data from existing hydroelectric dams to get empirical data and validation.

If we find that our model is inadequate or that it fails in some way, we then enter an iterative loop in which we cycle back to an earlier stage of the model building and re-examine our assumptions, our known parameter values, the principles chosen, the equations used, themeans of calculation, and so on. This iterative process is essential because it is the only way that models can be improved, corrected, and validated.

1.3 Some Methods of Mathematical Modeling

Now we will review some of the mathematical techniques we can use to help answer the philosophical questions posed in Section 1.2. These mathematical principles include: dimensional homogeneity, abstraction and scaling, conservation and balance principles, and consequences of linearity. We will expand these themes more extensively in the first part of this book.

1.3.1 Dimensional Homogeneity and Consistency

There is a basic, yet very powerful idea that is central to mathematical modeling, namely, that every equation we use must be dimensionally homogeneous or dimensionally consistent. It is quite logical that every term in an energy equation has total dimensions of energy, and that every term in a balance of mass should have the dimensions of mass. This statement provides the basis for a technique called dimensional analysis that we will discuss in greater detail in Lecture 2.

In that discussion we will also review the important distinction between physical dimensions that relate a (derived) quantity to fundamental physical quantities and units that are numerical expressions of a quantity's dimensions expressed in terms of a given physical standard.

1.3.2 Abstraction and Scaling

An important decision in modeling is choosing an appropriate level of detail for the problem at hand, and thus knowing what level of detail is prescribed for the attendant model. This process is called abstraction and it typically requires a thoughtful approach to identifying those phenomena on which we want to focus, that is, to answering the fundamental question about why a model is being sought or developed.

For example, a linear elastic spring can be used to model more than just the relation between force and relative extension of a simple coiled spring, as in an old-fashioned butcher's scale or an automobile spring. It can also be used to model the static and dynamic behavior of a tall building, perhaps to model wind loading, perhaps as part of analyzing how the building would respond to an earthquake. In these examples, we can use a very abstract model by subsuming various details within the parameters of that model.

As we talk about finding the right level of abstraction or the right level of detail, we are simultaneously talking about finding the right scale for the model we are developing. For example, the spring can be used at a much smaller, micro scale to model atomic bonds, in contrast with the macro level for buildings. The notion of scaling includes several ideas, including the effects of geometry on scale, the relationship of function to scale, and the role of size in determining limits—all of which are needed to choose the right scale for a model in relation to the "reality" we want to capture.

1.3.3 Conservation and Balance Principles

When we develop mathematical models, we often start with statements that indicate that some property of an object or system is being conserved. For example, we could analyze the motion of a body moving on an ideal, frictionless path by noting that its energy is conserved. Sometimes, as when wemodel the population of an animal colony or the volume of a river flow, we must balance quantities, of individual animals or water volumes, that cross a defined boundary. We will apply balance or conservation principles to assess the effect of maintaining or conserving levels of important physical properties. Conservation and balance equations are related—in fact, conservation laws are special cases of balance laws.

The mathematics of balance and conservation laws are straightforward at this level of abstraction. Denoting the physical property beingmonitored as $Q(t)$ and the independent variable time as t , we can write a balance law for the temporal or time rate of change of that property within the system boundary depicted in Figure 1.3 as:

$$\frac{dQ(t)}{dt} = q_{in}(t) + g(t) - q_{out}(t) - c(t) \qquad (1.1)$$

where $q_{in}(t)$ and $q_{out}(t)$ represent the flow rates of Q(t) into (the *influx*) and out of (the *efflux*) the system boundary, g (t) is the rate at which Q is generated within the boundary, and c(t) is the rate at which Q is consumed within that boundary. Note that eq. (1.1) is also called a *rate equation* because each term has both the meaning and dimensions of the rate of change with time of the quantity Q(t).

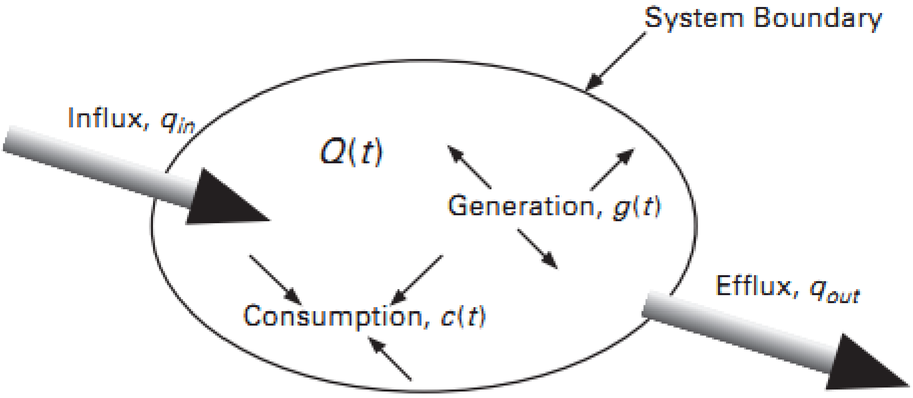

Figure 1.3 A system boundary surrounding the object or system being modeled. The influx $q_{in}(t)$, efflux $q_{out}(t)$, generation $g(t)$, and consumption $c(t)$, affect the rate at which the property of interest, $Q(t)$, accumulates within the boundary (after Cha, Rosenberg, and Dym, 2000).

In those cases where there is no generation and no consumption within the system boundary (i.e., when g = c = 0), the balance law in eq. (1.1) becomes a conservation law:

$$\frac{dQ(t)}{dt} = q_{in}(t) - q_{out}(t) \qquad (1.2)$$

Here, then, the rate at which Q(t) accumulates within the boundary is equal to the difference between the influx, $q_{in}(t)$, and the efflux, $q_{out}(t)$.

1.3.4 Constructing Linear Models

Linearity is one of the most important concepts in mathematical modeling. Models of devices or systems are said to be linear when their basic equations—whether algebraic, differential, or integral—are such that the magnitude of their behavior or response produced is directly proportional to the excitation or input that drives them. Even when devices like the pendulum discussed in Chapter 7 are more fully described by nonlinear models, their behavior can often be approximated by linearized or perturbed models, in which cases the mathematics of linear systems can be successfully applied.

We apply linearity when we model the behavior of a device or system that is forced or pushed by a complex set of inputs or excitations. We obtain the response of that device or system to the sum of the individual inputs by adding or superposing the separate responses of the system to each individual input. This important result is called the principle of superposition. Engineers use this principle to predict the response of a system to a complicated input by decomposing or breaking down that input into a set of simpler inputs that produce known system responses or behaviors.

1.4 Model classifications in mathematics

Mathematical models are usually composed of relationships and variables. Relationships can be described by operators, such as algebraic operators, functions, differential operators, etc. Variables are abstractions of system parameters of interest, that can be quantified. Several classification critera can be used for mathematical models according to their structure:

Linear vs. nonlinear: If all the operators in a mathematical model exhibit linearity, the resulting mathematical model is defined as linear. A model is considered to be nonlinear otherwise. The definition of linearity and nonlinearity is dependent on context, and linear models may have nonlinear expressions in them. For example, in a statistical linear model, it is assumed that a relationship is linear in the parameters, but it may be nonlinear in the predictor variables. Similarly, a differential equation is said to be linear if it can be written with linear differential operators, but it can still have nonlinear expressions in it. In a mathematical programming model, if the objective functions and constraints are represented entirely by linear equations, then the model is regarded as a linear model. If one or more of the objective functions or constraints are represented with a nonlinear equation, then the model is known as a nonlinear model.

Nonlinearity, even in fairly simple systems, is often associated with phenomena such as chaos and irreversibility. Although there are exceptions, nonlinear systems and models tend to be more difficult to study than linear ones. A common approach to nonlinear problems is linearization, but this can be problematic if one is trying to study aspects such as irreversibility, which are strongly tied to nonlinearity.

Static vs. dynamic: A dynamic model accounts for time-dependent changes in the state of the system, while a static (or steady-state) model calculates the system in equilibrium, and thus is time-invariant. Dynamic models typically are represented by differential equations.

Explicit vs. implicit: If all of the input parameters of the overall model are known, and the output parameters can be calculated by a finite series of computations (known as linear programming, not to be confused with linearity as described above), the model is said to be explicit. But sometimes it is the output parameters which are known, and the corresponding inputs must be solved for by an iterative procedure, such as Newton's method (if the model is linear) or Broyden's method (if non-linear). For example, a jet engine's physical properties such as turbine and nozzle throat areas can be explicitly calculated given a design thermodynamic cycle (air and fuel flow rates, pressures, and temperatures) at a specific flight condition and power setting, but the engine's operating cycles at other flight conditions and power settings cannot be explicitly calculated from the constant physical properties.

Discrete vs. continuous: A discrete model treats objects as discrete, such as the particles in a molecular model or the states in a statistical model; while a continuous model represents the objects in a continuous manner, such as the velocity field of fluid in pipe flows, temperatures and stresses in a solid, and electric field that applies continuously over the entire model due to a point charge.

Deterministic vs. probabilistic (stochastic): A deterministic model is one in which every set of variable states is uniquely determined by parameters in the model and by sets of previous states of these variables; therefore, a deterministic model always performs the same way for a given set of initial conditions. Conversely, in a stochastic model—usually called a "statistical model"—randomness is present, and variable states are not described by unique values, but rather by probability distributions.

Deductive, inductive, or floating: A deductive model is a logical structure based on a theory. An inductive model arises from empirical findings and generalization from them. The floating model rests on neither theory nor observation, but is merely the invocation of expected structure. Application of mathematics in social sciences outside of economics has been criticized for unfounded models. Application of catastrophe theory in science has been characterized as a floating model.

1.5 Summary

In this chapter we have provided an overview of the foundational material we will cover in this book. In so doing, we have defined mathematical modeling, provided motivation for its use in engineering and science, and set out a principled approach to doing mathematical modeling. We have also outlined some of the important tools that will be covered in greater detail later: dimensional analysis, abstraction and scaling, balance laws, and linearity.

It is most important to remember that mathematical models are representations or descriptions of reality—by their very nature they depict reality. Thus, we close with a quote froma noted linguist (and former senator from

California) to remind ourselves that we are dealing with models that, we hope, represent something that seems real and relevant to us. However, they are abstractions and models, they are themselves real only as models, and they should never be confused with the reality we are trying to model. Thus, if the behavior predicted by our models does not reflect what we see or measure in the real world, it is the models that need to be fixed—and not the world:

"The symbol is NOT the thing symbolized; the word is NOT the thing; the map is NOT the territory it stands for."

—S. I. Hayakawa, *Language in Thought and Action*

A mathematical model is a description of a system using mathematical concepts and language. The process of developing a mathematical model is termed mathematical modeling. Mathematical models are used in the natural sciences (such as physics, biology, earth science, meteorology) and engineering disciplines (such as computer science, artificial intelligence), as well as in the social sciences (such as economics, psychology, sociology, political science). Physicists, engineers, statisticians, operations research analysts, and economists use mathematical models most extensively. A model may help to explain a system and to study the effects of different components, and to make predictions about behaviour.

Mathematical models can take many forms, including but not limited to dynamical systems, statistical models, differential equations, or game theoretic models. These and other types of models can overlap, with a given model involving a variety of abstract structures. In general, mathematical models may include logical models. In many cases, the quality of a scientific field depends on how well the mathematical models developed on the theoretical side agree with results of repeatable experiments. Lack of agreement between theoretical mathematical models and experimental measurements often leads to important advances as better theories are developed.