1

INTRODUCTION

Molecular dynamics simulations compute the motions of individual molecules in models of solids, liquids, and gases. The key idea here is *motion*, which describes how positions, velocities, and orientations change with time. In effect, molecular dynamics constitutes a motion picture that follows molecules as they dart to and fro, twisting, turning, colliding with one another, and, perhaps, colliding with their container.

This usage is not unique: *molecular dynamics* may also refer to the motions of real molecules when studied primarily by molecular beam [1] or spectroscopic [2] techniques. This terminological confusion is compounded by *lattice dynamics* [3], which refers to the study of vibratory motions of atoms in solids, and by *molecular mechanics* [4], also called force field calculations, which refers to quantum mechanical calculations of the structure of individual molecules. This book is concerned with molecular dynamics solely in the sense of simulation.

Molecular dynamics simulation is the modern realization of an old, essentially old-fashioned, idea in science; namely, the behavior of a system can be computed if we have, for the system's parts, a set of initial conditions plus forces of interaction. From the time of Newton to the present day, this deterministic mechanical interpretation of Nature has dominated science [5]. In 1814, roughly a century after Newton, Laplace wrote [6]:

> Given for one instant an intelligence which could comprehend all the forces by which nature is animated and the respective situation of the beings who compose it—an intelligence sufficiently vast to submit these data to analysis—it would embrace in the same formula the movements of the greatest bodies of

> the universe and those of the lightest atoms; for it, nothing would be uncertain and the future, as the past, would be present to its eyes.

If this approach is thwarted by the complexities of reality, then we replace reality with a model. In one of his Baltimore lectures (Lecture XI), roughly a century after Laplace, Thomson observed [7]:

> It seems to me that the test of "Do we or not understand a particular subject in physics?" is, "Can we make a mechanical model of it?"

Today, roughly a century after Thomson, we remain undeterred from Laplace's dream: the requisite "intelligence" is provided by the digital computer, the "respective situation" is a set of initial positions and velocities, "the same formula" though not literally true could be interpreted as the same algorithmic program, and Laplace's universe has given way to model universes. Now, deterministic mathematical models pervade not only the physical sciences and engineering, but the life and social sciences [8] as well.†

This attitude is old-fashioned in the sense that, while often successful, it is nevertheless simplistic. In spite of Laplace's claim, we can still identify systems that are unpredictable—stock markets and the weather, for example. Why should this be? If deterministic mathematical models can help us successfully land *Apollo XI* on the moon, why can't they help us predict next month's weather on earth?

The resolution of this dilemma is based on the kind of forces acting among system components: when a system contains objects that interact *nonlinearly*, the system's behavior may be unpredictable. In the past few years studies in nonlinear dynamics have decoupled deterministic from predictable [9]. *Deterministic* situations have system outputs causally connected to system inputs. *Calculable* situations are those deterministic situations in which an algorithm allows us to compute system outputs if the inputs were known. *Predictable* situations are those calculable situations in which the algorithm can be numerically implemented to actually compute the outputs. Calculable situations may be unpredictable because of the large number of inputs needed, because of an unrealistically high precision with which the inputs must be known, and/or because the algorithm's stability is sensitive to intermediate calculations. In pool, *Eight ball in the side pocket* is deterministic, calculable, and predictable; however, whether it will rain in two weeks is deterministic but unpredictable.

The overriding theme of this book is predicated on the decoupling of predictability from determinism. Be warned—that you use a machine to compute the behavior of a many-body model does not guarantee that the computed behavior is representative of that model, much less that the model mimics reality. To my mind it is this deterministic unpredictability that makes molecular dynamics fascinating and challenging. Is the fun (aka intellectual stimulation) merely in making a model, writing some differential equations, and loading it all into a computer? No. The fun, it seems to me, begins when we have completed a simulation, when we have a number. Now there arise all the old familiar questions characteristic of science: How good is this number? How could it be wrong? Is it representative, that is, reproducible? What does it mean? Do I believe it? How do I test it? If it is right, what must follow? This book should not only help you learn how to simulate but also make you aware that questioning the results is part of the procedure.

Computer simulations are performed on models, not on real things, and so the science of simulation, while distinct from, is necessarily bound to the art of model building. The purposes of this chapter are to clarify the distinctions between models and simulations and to discuss how together they contribute to new understanding.

1.1 SYSTEMS AND ALL *THAT*

The portion of the physical world on which we focus our attention is called the *system*; it is a subset of the universe. The system may be composed of any number of similar or dissimilar parts and the condition of those parts identifies the *state* of the system. For example, the door to my office constitutes a system to which I can ascribe two states: *open* or *shut*. To analyze and describe the behavior of the system, we need ways for assigning numerical values either to the state or to functions of the state; such assignments are called *observables*. Thus, to my door we can ascribe an observable called openness, to which I assign the value 1 if the door is open and 0 if the door is shut. As another example, let the system be 10^{24} molecules of a gas. Its state is specified by the position and momentum of each molecule, and the state gives rise to such observables as temperature and pressure.

The state of a system can be manipulated and controlled from the environment via *interactions*. For example, I may change the state of my office door by an interaction, specifically, by exerting a force. Allowed interactions are constrained by the nature of the boundary that separates the system from its surroundings. Various kinds of boundaries and interactions are possible, but in this book we will limit our attention to *isolated systems*: systems that can exchange neither matter nor energy with their surroundings.

We cannot usually study a system by directly observing the state; instead, we probe states indirectly by manipulating, controlling, and measuring observables. Thus, studies of a gas may involve controlling the system volume, manipulating its temperature, and measuring how the pressure responds. The

---

†In fairness, paleontologists, at least, have discovered deterministic unpredictability. Thus, Stephen Jay Gould [10] posits that if the tape of life were rewound to some previous, sufficiently removed condition and then replayed, the result would be life unlike life as we know it. For more technical conjectures on connections between life and deterministic unpredictability, see Fox [11].

isolated system is special in the sense that we do not interact with it, and therefore we can manipulate its observables only before the system is isolated.

To organize, describe, and perhaps even predict observables, we create theories. Theories may operate at one of several levels. At the simplest level, theories merely provide relations among observables. For example, the ideal-gas law

$$PV = NkT \tag{1.1}$$

was originally obtained by organizing results from measurements of the pressure, volume, and temperature of low-density gases.

At the next level of complexity are theories that relate observables to the underlying state. For example, at this level we have kinetic theory, which teaches that the observable temperature is related to the state through the molecular velocities. Theories at this level provide interpretations or explanations for observables, but if the state itself is unobservable, these theories cannot be used to compute numerical values for observables.

To overcome this computational dead end, two strategies have been devised: (1) concoct theories at still higher levels or (2) perform computer simulations. Higher level theories try to resolve the computational difficulty by reorganizing and reducing the detailed information about the state needed to compute values for observables. Such is the objective of statistical mechanics, in which observables are related not to the underlying state itself, but rather to the probability of the system being in particular states.

The alternative strategy includes molecular dynamics. Molecular dynamics assigns numerical values to states, thereby making states observable, at least for model substances. With numerical values assigned to states, theoretical relations from kinetic theory can be used to compute values for experimentally accessible observables. Thus molecular dynamics is closely tied to kinetic theory and not as closely related to statistical mechanics.[†] In particular, molecular dynamics is less sophisticated, less elegant, but more direct than statistical mechanics.

1.2 MODELING VERSUS SIMULATION

Whether we study systems theoretically or experimentally, the general procedure is the same: we manipulate and control certain observables (inputs), the

system responds, and then we measure or compute other observables (outputs):

$$\begin{array}{c}\text{Manipulate and control}\\ \text{certain observables}\\ \text{(inputs)}\end{array} \rightarrow \boxed{\text{SYSTEM}} \rightarrow \begin{array}{c}\text{measure or compute}\\ \text{other observables}\\ \text{(outputs)}\end{array}$$

Since theoretical analyses are now largely done via modeling or simulation or both, it is instructive to clarify how modeling differs from simulation. For a discussion more complete than what follows, see Casti [8].

The goal of theoretical work is to establish connections between measurable outputs and controlled inputs. In Section 1.1 we discussed how this may be done at different levels of complexity; in particular, sophisticated theories use an underlying state to connect outputs to inputs. Part of the theoretical problem is to define the state in such a way that complicated interactions among state variables are decoupled, or at least weakened, so that observable outputs can in fact be computed.

A model is an attempt to decouple and remove interactions that have little or no influence on the observables being studied. Thus, a model is simpler than the system it mimics: it has access to fewer states. Decoupling interactions means relaxing constraints; hence, a model has access to some states not available to the original system and vice versa. In other words, a model is a subset or subsystem of the original system: outputs from a model will be consistent with those of the original system, but only for a restricted set of inputs. For those restricted inputs, since the model is a subsystem of the original, states visited by the model correspond to those visited by the original system.

In contrast, a simulation is more complicated than the system it simulates: a simulation generally can reach many more states than can the original system. A simulation imposes constraints so that the simulated output is consistent with the output of the original system, at least for a restricted set of inputs. A simulation will typically bear no structural relation to the original system; for example, the way constraints are imposed in the simulation may differ from the mechanism that confines the original system to certain states. Hence, states in the simulation may bear no correspondence to states of the original system. Although a simulation is more complex than the original system, it does *not* follow that the original system is a model of the simulation.

An example should clarify these ideas. As the real system, consider a simple ball-and-spring arrangement. One end of the spring is attached to the ball, the other end is fixed to a wall. In response to a displacement from its equilibrium position, the ball slides on a floor. The problems are to, in turn, model and simulate the motion of the ball that results from a displacement.

[†] W. Thomson, Lecture I, p. 1: "…the kinetic theory of gases is a part of molecular dynamics, is founded upon molecular dynamics, works wholly within molecular dynamics, to it molecular dynamics is everything, and it must be advanced by molecular dynamics…"

Let R represent this real system, that is, the spring as mover plus the ball as the thing moved.

To study the motion of R, we might construct a device M that is a one-dimensional harmonic oscillator (ODHO) having spring constant γ. The motion of M is described by the differential equation

$$\frac{d^2x}{dt^2} = -\gamma x(t) \tag{1.2}$$

where x is the distance of the ball from its equilibrium position and t is time. Writing this equation presumes several simplifying assumptions: (a) the motion of the ball is restricted to a line, (b) the spring is perfectly harmonic, (c) the ball experiences no sliding friction on the floor, and (d) the ball has no internal states that exchange energy with the spring. These assumptions imply that M is a simple subsystem of R. For some initial displacements M will mimic the motion in R; however, M cannot mimic all the behavior available to R. For example, in R we might initially raise the ball from the floor, allowing the ball to move in the xz-plane, motion not allowed to M. The one-dimensional harmonic oscillator M is a *model* of the real system R.

An alternative scheme S for studying the motion of R would be to remove the spring and use a person as the mover. This person might be a well-trained graduate student who has the uncanny ability to move the ball, for many sets of initial conditions, so as to reproduce the motion of the ball in R. This situation S is more complex than R because a person is more sophisticated than a spring. Moreover, S involves imposing constraints on the student's arm to make the motion of the ball mimic the motion in R. Otherwise, more states are possible in S than in R; for example, the student might absent-mindedly drop the ball in his pocket when he stops for coffee. The situation S is a *simulation* of the real system R.

Note that as well as simulating the real system R, we could also simulate the model M. For example, we might have another mover (a professor or, equivalently, a robot) that can move the ball through states visited by the perfect ODHO model M. This simulation of M is more complex than M itself, but the added complexity does not make the model more realistic, that is, more like the real system R.

How do these distinctions relate to molecular simulations? Well, what do we typically do? We identify a substance and its observables that we want to study, say, thermodynamic properties of argon. Then we construct a model of the substance, say, the spherically symmetric, pairwise additive Lennard-Jones potential. This is a true model. The Lennard-Jones potential is simpler than the argon potential because argon atoms are not perfect spheres and their interactions are certainly not only pairwise additive. With the model chosen, we then perform a simulation—but a simulation of what? It can only be a simulation of the model, of the Lennard-Jones substance. We do not simulate argon. The simulation is more complex than the model, but the added complexity does not add to the realism of the resulting observable outputs. In error are those who claim that molecular dynamics simulates argon, or water, or proteins, or whatever. We simulate molecular models of such substances.

1.3 THEORY VERSUS EXPERIMENT

At a scientific conference in the 1970s there broke out a heated debate as to whether computer simulations like molecular dynamics are theories or are experiments. The theory side argued that simulation is clearly not experiment because no measurements are done on real systems; molecular simulations are pure calculation. The experiment side countered that simulation results are *used* like experiments, namely, to test theories; it isn't sensible to test one theory with another theory is it? Moreover, this side noted that simulation results, like experiments, are prone to problems of reproducibility and statistical error. Hence, pervading the literature is the interpretation of molecular simulations as computer experiments.

What's the resolution of this dilemma? And does it really matter how we think of simulation? Consider the following example. To perform an experiment, to take a measurement, the observer must interact with the system: some type of probe necessarily has to cross the system boundary. Thus, truly isolated systems cannot be studied experimentally: once our probe crosses the boundary, the system is no longer isolated. However, we can perform theoretical calculations, such as simulations, on truly isolated systems and obtain meaningful results. The resolution of our dilemma has to be that molecular simulations are forms of theory; they do not involve measurements on real systems.

How we think of simulation is important, indeed crucial, because of the consequences: if we accept that simulations are experiments, then it follows that the models simulated are real. Armed with this attitude plus the ease of actually doing simulations, we may be tempted to abandon laboratory experiment altogether. The danger lies in severing simulation from reality.

Is this only an academic issue? Recently a public presentation was made by the chief executive officer of a major computer manufacturer. In the talk *Voyager* photos of Jupiter were compared with images produced from a computer simulation of Jovian weather. The speaker's punch line was that the simulated images were actually "more real than real life." But if this were true, then there would be no need for further explorations by spacecraft —we would simply perform simulations. Note that, in fact, reality doesn't enter this picture at all. The CEO can be faulted on two counts: not only did he erroneously claim that a simulation can supplant reality, but he also confused a photographic image with reality.[†]

To illustrate this point in a more mundane situation, consider: Police have been summoned to the scene of a domestic quarrel. In the kitchen the patrolmen find a bewildered wife standing over the prostrate form of her husband.

> "Okay, lady, what happened?"
>
> "I don't rightly know. Herb riled me, so I hit'em in the head with a tomato."
>
> "Sure, lady. There's no tomato on the body or the floor."
>
> "Isn't there? O'course not. T'was a decorative ceramic."

And lest you feel that this contrived example[‡] is only impractical philosophical quibbling, consider the intense ethical debates that were prompted by American television broadcasts of selected news events in the form of interpolative reenactments (aka *simulations*).

1.4 REDUCTIONISM VERSUS SIMULATION

Since the time of Newton, scientific theories have nearly all been developed in a reductionistic mode: a complex system is reduced to one or more simple subsytems and the subsystems are analyzed. Subsystems ultimately take the form of models, and today models are almost exclusively mathematical. Before about 1960 mathematical models had to be simple enough to be tractable analytically, but now this constraint is relaxed by the availability of digital computers.

As an example of reductionism, consider the study of matter in simpler and simpler forms:

Matter → compounds → elements → molecules → atoms

→ elementary particles → quarks

Modeling may occur at any stage of such a reduction. Successful modeling requires a construction that forces the behavior of interest to remain *invariant* when the subsystem is replaced by the model. The goal of reductionism is to explain system behavior by combining explanations for the behaviors of its subsystem models. Simulation provides an alternative to reductionism because simulation allows us to study the behaviors of classes of systems or subsystems. Thus, while reductionism emphasizes structural analysis, simulation emphasizes behavioral classification. As an example, consider study of the fluid–solid phase transition. One simulation approach to this problem would be to load marbles into a drum. To force the marbles to move, we rotate the drum (imagine a cement mixer). As the drum rotates, we add more and more marbles, until finally enough are added to freeze the motion. From the number of marbles N and the drum volume V, we obtain the density N/V needed for solidification. We then repeat the simulation with balls of other diameters: golf balls, baseballs, soccer balls, basketballs. By inspecting the solidification densities for balls of various diameters, we conclude that one mode of solidification is a geometric packing effect, controlled by the packing fraction $V_{\text{balls}}/V_{\text{drum}}$. Note the features of this simulation: the controlled input observables are the ball diameter, the number of balls, and the drum volume; the measured output observable is the density at which motion ceases. During the experiments, the states of the balls—positions and velocities—bear no relation to states of molecules in any real substance. Further, the entire study is not just of one substance, but rather a systematic progression through a class of substances: spheres of increasing diameters.

In contrast, a reductionist would study this problem by combining a model with a theory to predict the solidification of a particular substance. The theory might involve only relations among observables, such as PVT equations of state, or the theory might include underlying system states, such as is done in statistical mechanics. In any case, the validity of the model and the theory would be tested by comparing predictions with experimental measurements on a real substance. However, the connection between input and output observables would remain *implicit* in the mathematical apparatus used to make the prediction.

But the goal of the simulations is not so much to predict solidification as to make *explicit* how input and output observables are connected. In other words, rather than predictions, the goal is more in the nature of providing explanations: idealized models "*explain* nature even while they do not describe it" [14]. Both reductionism and simulation contribute to science; however, for a specific problem one or the other may be more appropriate. In particular, simulation is not always the best method.

As shown in Figure 1.1, we identify two distinct roles that simulation can play in scientific investigation. At the higher level, simulation, including computer simulation, serves as an alternative to reductionism. At this level, as popularly claimed, simulation is a new way of doing science. In addition, computer simulation can be used at a lower level, as a tool in reductionism. It is this second use of simulation that is implied by the more familiar triangular diagram shown in Figure 1.2 [15]. That diagram suggests two reductionistic uses for simulation: (a) simulation data on models can be used to test

---

[†] A principal feature of science is the apparently endless disentangling of images from reality. Well over 800 years ago in his *Questiones Naturales*, the scholar Adelard of Bath was moved to write, "Wherefore, if you want to hear anything more from me, give and take reason. For I'm not the sort of man that can be fed on a picture of a beefsteak" [12].

[‡] This example is a slight modification of a vaudeville routine used by James Thurber [13] to illustrate *confusing the container with the thing contained*.

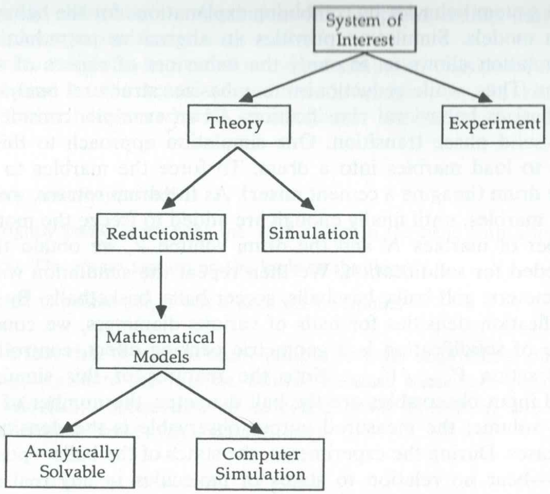

FIGURE 1.1 Hierarchy of scientific modes of investigation. Note that the system of interest may be real, or it may itself be a model.

theories and (b) simulation data can be compared with experimental data to test the realism of simulated models.

In Figure 1.1, the dual use—as an alternative to reductionism or as a reductionistic tool—explains, perhaps, early debates over whether computer simulation is theory or experiment. If you identify all theory as reductionism and sense that simulation is something different, then you may interpret simulation as experiment. Conversely, if you see computer simulation as a tool for studying reductionistic models, then you may interpret simulation as theory.

The theme of this section is that computer simulation offers possibilities more instructive and more far-reaching if it is used as an alternative to reductionism rather than as merely a servant to reductionism. To make this statement concrete: using computer simulation to map out a phase diagram for a model of methane (or whatever), merely to test how realistic the model may be, is to misunderstand and underutilize the power of simulation.

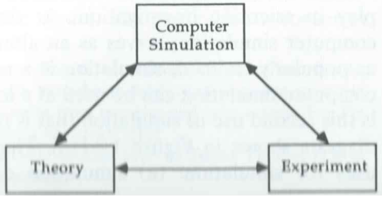

FIGURE 1.2 Conventional representation of the interplay among theory, experiment, and computer simulation.

1.5 MODELS FOR MOLECULAR SIMULATIONS

A computer simulation is valuable because it is applied to a precisely defined model for the material of interest. The model is actually a composite of two: one for interactions among the molecules making up the system and another for interactions between the molecules and their environment:

$$\boxed{\text{Simulated model}} = \boxed{\text{model for molecular interactions}}$$

$$+ \boxed{\text{model for system–environment interactions}}$$

Note the decoupling implied by this schematic—intermolecular interactions are presumed to be independent of interactions with the environment.

The model for molecular interactions is contained in an intermolecular force law or, equivalently, an intermolecular potential energy function. This potential function implicitly describes the geometric shapes of individual molecules or, more precisely, their electron clouds. Thus when we specify the potential function, we establish the symmetry of the molecules, whether they are rigid or flexible, how many interaction sites occupy each molecule, and so on. A detailed characterization of intermolecular potential functions may be given analytically or numerically; in any case, a quantitative form for the potential function defines a molecular model and hence the form must be chosen before a simulation can be performed.

In this book we consider only spherically symmetric molecules (atoms). For N such atoms the intermolecular potential function is represented by $\mathscr{U}(\mathbf{r}^N)$. The notation $\mathbf{r}^N$ represents the set of vectors that locate the atomic centers of mass, $\mathbf{r}^N = \{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \ldots, \mathbf{r}_N\}$. When we establish values for the set $\mathbf{r}^N$, we define the *configuration* of a system. Macroscopic properties that are averages over only the set $\mathbf{r}^N$ are called *configurational* properties.

In most simulations the intermolecular potential energy is taken to be a sum of isolated pair interactions; this assumption is called *pairwise additivity*. Hence,

$$\mathscr{U} = \sum_{i<j}\sum u(r_{ij}) \qquad (1.3)$$

where $u(r_{ij})$ is a pair potential energy function whose form is known and r_{ij} is the scalar distance between molecules i and j. Since no dissipative forces

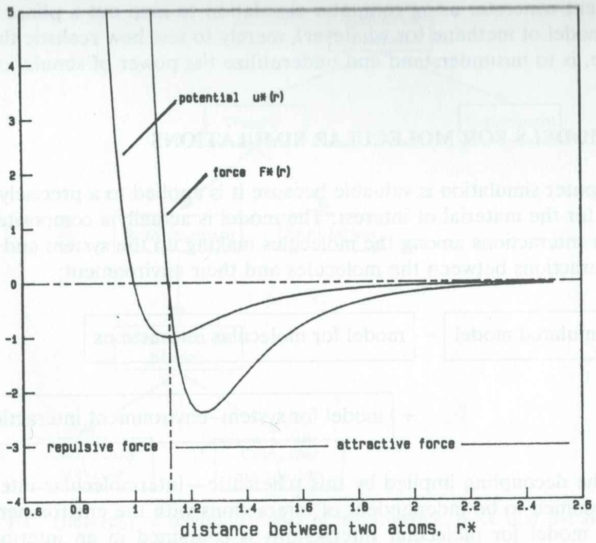

FIGURE 1.3 Typical potential energy $u^*(r)$ and force $F^*(r)$ functions used to model interactions between two atoms separated by distances r^*. Positive forces are repulsive; negative forces attractive. Here $r^* = r/\sigma$, $u^* = u/\varepsilon$, and $F^* = F\sigma/\varepsilon$, where σ is a unit of length and ε is a unit of energy.

act among molecules, intermolecular forces are conservative, and therefore the force on molecule i is related to the potential by

$$\mathbf{F}_i = -\frac{\partial \mathscr{U}(\mathbf{r}^N)}{\partial \mathbf{r}_i} \tag{1.4}$$

where $\partial/\partial\mathbf{r}$ represents the gradient operator. Typical forms for the pair potential and its resultant force are shown in Figure 1.3. By convention, repulsive forces are positive, attractive forces are negative.

The second part of the simulated model encompasses boundary conditions, which describe how the molecules interact with their surroundings. Characteristics of boundary conditions are largely dictated by the physical situation to be simulated; however, some freedom usually exists in the way in which boundary conditions are realized. For example, if a bulk fluid is to be studied, we want to avoid hard boundaries, which obscure bulk–fluid behavior in small systems. If nonuniform regions such as fluid–fluid or fluid–solid interfaces are to be simulated, then boundary conditions that mimic these situations are needed. If the shear or bulk viscosity is to be determined, we may introduce moving boundaries that shear or compress the system. In this book we consider only boundaries that isolate the system, so no interactions occur between system and surroundings. In any event, setting boundary conditions completes the definition of the model system to be simulated.

1.6 STOCHASTIC VERSUS DETERMINISTIC

A molecular-scale simulation consists of three principal steps: (1) construction of a model, (2) calculation of molecular trajectories, and (3) analysis of those trajectories to obtain property values. The second step constitutes the simulation proper. We use the way in which molecular positions $\mathbf{r}^N$ are computed in step 2 to discriminate among simulation methods.

In molecular dynamics the positions are obtained by numerically solving differential equations of motion and, hence, the positions are connected in time—the positions reveal dynamics of individual molecules as in a motion picture. In other simulation methods the molecular positions are not temporally related. For instance, in *Monte Carlo* simulations the positions are generated stochastically such that a molecular configuration $\mathbf{r}^N$ depends only on the previous configuration. When the outcome of a random event in a sequence depends only on the outcome of the immediately previous event, the sequence is called a *Markov chain*. In still other simulation methods, the positions are computed from hybrid schemes that involve some stochastic features, as in Monte Carlo, and some deterministic features, as in molecular dynamics. These various methods can be arranged, as in Figure 1.4, according to the degree of determinism used in generating molecular positions. Here some of these methods are briefly described to place molecular dynamics within the wider context of molecular-scale simulation methods.

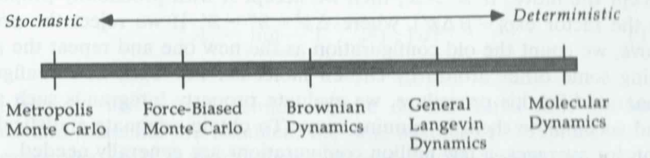

FIGURE 1.4 Relative degree of determinism in various molecular simulation methods. (Adapted from Ceperley and Tully [16] and used with permission.)

1.6.1 Monte Carlo

The purely stochastic method, Monte Carlo, is typically performed on a fixed number of molecules N placed in a fixed volume V and maintained at a constant absolute temperature T. The simulation procedure is adapted from general Monte Carlo methods for evaluating multidimensional integrals. Here the integrals of interest are statistical–mechanical ensemble averages, that is, configurational properties $\langle A \rangle$ of the N-body system. For atomic substances these integrals are of the form

$$\langle A \rangle = \frac{1}{\mathscr{Z}} \int \cdots \int \exp[-\beta \mathscr{U}(\mathbf{r}^N)] A(\mathbf{r}^N)\, d\mathbf{r}_1 \cdots d\mathbf{r}_N \tag{1.5}$$

where $\beta = 1/kT$, k is Boltzmann's constant, and $\mathscr{Z}$ is the configurational integral

$$\mathscr{Z} = \int \cdots \int \exp[-\beta \mathscr{U}(\mathbf{r}^N)]\, d\mathbf{r}_1 \cdots d\mathbf{r}_N \tag{1.6}$$

The integrals in (1.5) and (1.6) are $3N$-fold because each differential volume element contains three components; thus, $d\mathbf{r}_1 = dx_1\, dy_1\, dz_1$.

In Monte Carlo simulations, ensemble averages such as (1.5) are evaluated by accumulating the integrand at randomly generated values of the independent variables—the atomic positions $\mathbf{r}^N$. Note that because of the "Boltzmann" factor $\exp(-\beta\mathscr{U})$, some configurations make large contributions to the integral while others contribute nothing. Thus we seek to bias the sampling in favor of those configurations most likely to occur. The importance sampling scheme devised by Metropolis et al. [17] does just that.

The Metropolis method involves the following principal steps. First, we assign initial positions $\mathbf{r}_i$ to the N molecules and compute the total potential energy $\mathscr{U}$ using (1.3). Then we hypothesize a new configuration by arbitrarily choosing one molecule and proposing that it be moved from $\mathbf{r}$ through a randomly chosen distance and direction to a new position $\mathbf{r}'$. We compute the total potential energy $\mathscr{U}'$ for this new configuration and accept the proposed move (allow it to occur) based on the following criteria. If $\mathscr{U}' < \mathscr{U}$, then we accept the move. If $\mathscr{U}' > \mathscr{U}$, then we accept it with probability proportional to the factor $\exp[-\beta\,\Delta\mathscr{U}]$, where $\Delta\mathscr{U} = \mathscr{U}' - \mathscr{U}$. If we reject the proposed move, we count the old configuration as the new one and repeat the process using some other arbitrarily chosen molecule. For each new configuration generated by this procedure, we evaluate property integrands such as (1.5) and accumulate them in running sums. To obtain adequate statistical precision for averages, a few million configurations are generally needed.

Several variations on the Metropolis Monte Carlo method have been proposed. One of these is the force-biased algorithm [18] in which the proposed move of a molecule is no longer completely random; rather, its direction is taken to be that of the force exerted on the molecule by all other molecules. Such a procedure reduces the number of configurations needed for adequate statistical precision, but it increases the amount of computation per configuration [19].

A general presentation of Monte Carlo methods, including the Metropolis scheme, can be found in the book by Kalos and Whitlock [20]. The classic article on Monte Carlo applied to statistical mechanical problems is that by Wood [21]. More recent developments are contained in articles by Valleau and co-workers [22]. The books edited by Binder [23, 24] are good sources for applications of the method.

1.6.2 Molecular Dynamics

The molecular dynamics method encompasses two general forms: one for systems at equilibrium, another for systems away from equilibrium. As devised by Alder and Wainwright in the late 1950s [25], equilibrium molecular dynamics is typically applied to an isolated system containing a fixed number of molecules N in a fixed volume V. Because the system is isolated, the total energy E is also constant; here E is the sum of the molecular kinetic and potential energies. Thus the variables N, V, and E determine the thermodynamic state.

In NVE-molecular dynamics molecular positions $\mathbf{r}^N$ are obtained by solving Newton's equations of motion:

$$\mathbf{F}_i(t) = m\ddot{\mathbf{r}}_i(t) = -\frac{\partial \mathscr{U}(\mathbf{r}^N)}{\partial \mathbf{r}_i} \tag{1.7}$$

Here $\mathbf{F}_i$ is the force on i caused by the $N-1$ other molecules, the dots indicate total time derivatives, and m is the molecular mass. In writing (1.7) we have used (1.4), which relates the force to the intermolecular potential energy. Integrating (1.7) once yields the atomic momenta; integrating a second time produces the atomic positions. Repeatedly integrating for several thousand times produces individual atomic trajectories from which time averages $\langle A \rangle$ can be computed for macroscopic properties

$$\langle A \rangle = \lim_{t\to\infty} \frac{1}{t} \int_{t_0}^{t_0+t} A(\tau)\, d\tau \tag{1.8}$$

At equilibrium this average cannot depend on the initial time t_0. Since positions and momenta are obtained, the time average (1.8) represents both static properties, such as thermodynamics, and dynamic properties, such as transport coefficients.

According to the ergodic hypothesis, the time average (1.8) provided by molecular dynamics should be the same as the ensemble average (1.5) provided by Monte Carlo. Although a rigorous proof of the ergodic hypothesis exists only for the hard-sphere gas [26], it can be tested by comparing Monte Carlo results with molecular dynamics results. One such comparison is presented at the end of Chapter 3.

In Figure 1.5 the dynamic modeling problem is shown divided into its two great tasks: developing a suitable model for the problem at hand and applying molecular dynamics to that model. In this book we are not concerned with model development; instead, we concentrate on the second problem. That is, having developed a model, how do we perform the simulation? The simulation problem itself divides into two tasks: solving the equations of motion to generate molecular trajectories and then analyzing those trajectories for the properties of interest. In the figure we show how this book is organized to address those major tasks.

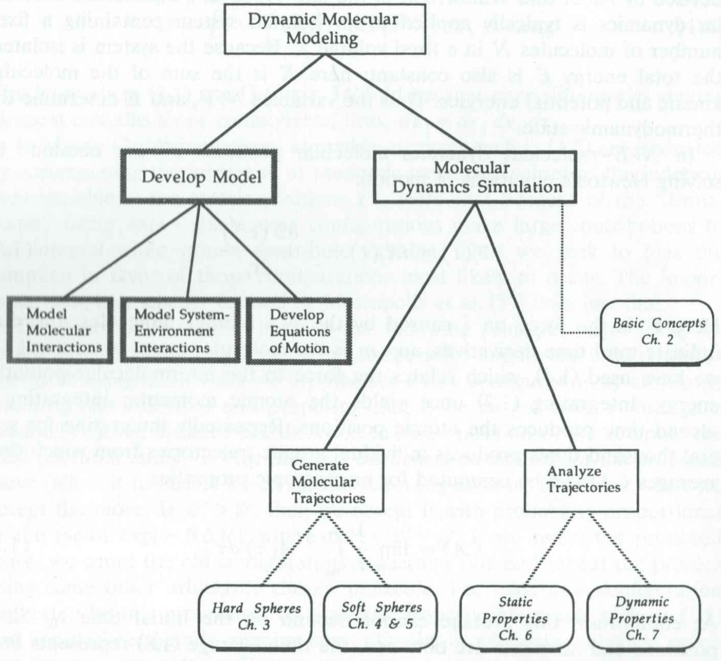

FIGURE 1.5 Hierarchy of the principal steps in dynamic molecular modeling. This book is not concerned with developing molecular models, but only with the application of molecular dynamics to existing models.

Molecular dynamics simulations are limited largely by the speed and storage constraints of available computers. Hence, simulations are usually done on systems containing 100–1000 particles, although calculations involving as many as 10^6 particles have been performed [27]. Because of this size limitation, simulations are confined to systems of particles that interact with relatively short-range forces; that is, intermolecular forces should be small when molecules are separated by a distance equal to half of the smallest overall dimension of the system. Because of the speed limitation, simulations are confined to studies of relatively short-lived phenomena, roughly, those occurring in less than 100–1000 psec. The characteristic relaxation time for the phenomenon under investigation must be small enough so that one simulation generates several relaxation times.

In addition to equilibrium molecular dynamics, nonequilibrium methods have been developed. These methods first appeared in the early 1970s [28–30], initially as alternatives to equilibrium simulations for computing transport coefficients. In these methods an external force is applied to the system to establish the nonequilibrium situation of interest, and the system's response to the force is then determined from the simulation. Nonequilibrium molecular dynamics has been used to obtain the shear viscosity, bulk viscosity, thermal conductivity, and diffusion coefficients [31].

Books devoted to molecular dynamics include that on simple fluids by Vesely [32] and those on nonequilibrium simulations by Hoover [33, 34] and by Evans and Morriss [35]. Simulations of plasmas, galactic evolution, and electron flow are emphasized in the book by Hockney and Eastwood [36]. The most comprehensive expositions of methods and applications are contained in the collection edited by Ciccotti and Hoover [37] and in the book by Allen and Tildesley [38].

1.6.3 Monte Carlo versus Molecular Dynamics

Although the physical and mathematical basis of Monte Carlo may be less transparent to a novice than that for molecular dynamics, Monte Carlo is usually easier than molecular dynamics to code in a high-level language such as Fortran. Monte Carlo is also easier to implement for systems in which it is difficult to extract the intermolecular force law from the potential function. Systems having this difficulty include those composed of molecules that interact through discontinuous forces; examples are the hard-sphere and hard convex-body models. Similar difficulties arise in systems for which the potential function is a complicated multidimensional surface, such as might be generated by ab initio calculations.

For determination of simple equilibrium properties such as the pressure in atomic fluids, Monte Carlo and molecular dynamics are equally effective: both require about the same amount of computer time to attain similar levels of statistical precision. However, molecular dynamics more efficiently evaluates such properties as the heat capacities, compressibilities, and interfacial properties. Besides configurational properties, molecular dynamics also provides access to dynamic quantities such as transport coefficients and time correlation functions. Such dynamic quantities cannot generally be obtained by Monte Carlo, although certain kinds of dynamic behavior may be deduced from Monte Carlo simulations [23].

Molecular dynamics also offers certain computational advantages because of the deterministic way in which it generates trajectories. The presence of an explicit time variable allows us to estimate the length needed for a run: the duration must be at least several multiples of the relaxation time for the slowest phenomenon being studied. No such convenient guide is available for estimating the length required for a Monte Carlo calculation. Finally, many kinds of small errors in a molecular dynamics program tend to accumulate with time and so become apparent as violations of conservation principles; in contrast, subtle errors in a Monte Carlo program may not blatantly advertise their presence.

1.6.4 Stochastic Dynamics

Molecular dynamics simulations of flexible molecules such as alkanes and proteins are more complex and time consuming than simulations of atomic and rigid-molecule systems. Not only are the equations of motion more complicated because of internal degrees of freedom, but also the internal modes, such as bond vibration and rotation, tend to relax on time scales very different from those of molecular collisions, which dominate external translational and rotational modes. Consequently, lengthy molecular dynamics simulations must be performed to capture many relaxation processes typical of flexible molecules.

If the objective is to study the dynamics of individual molecules, rather than determine properties of an aggregate, then Langevin and Brownian dynamics methods offer inexpensive alternatives to molecular dynamics. These methods reduce the amount of computation by modeling the system as one flexible molecule immersed in a continuum. For an interaction site i on a flexible molecule, the Langevin equation of motion is

$$m\ddot{\mathbf{r}}_i = -\zeta\dot{\mathbf{r}}_i + \mathbf{F}_{\text{intra}} + \mathbf{F}_{\text{inter}} \tag{1.9}$$

Here m is the mass of site i, ζ is a macroscopic friction coefficient, and $\mathbf{F}_{\text{intra}}$ is the force on i caused by interactions with other sites on the molecule. The intermolecular force $\mathbf{F}_{\text{inter}}$ is approximated as a stochastic interaction with the continuum.

In Brownian dynamics the momentum degrees of freedom are removed from (1.9) by arguing that over long times the positions of the interaction sites change very little. Thus setting the left-hand side of (1.9) to zero and solving explicitly for the positions $\mathbf{r}_i$, we obtain the Brownian equation of motion

$$\mathbf{r}_i(t) - \mathbf{r}_i(0) = \frac{1}{\zeta}\int_0^t [\mathbf{F}_{\text{intra}}(\tau) + \mathbf{F}_{\text{inter}}(\tau)]\, d\tau \tag{1.10}$$

Although (1.10) is approximate, it offers the advantage of allowing simulations of flexible molecules to be performed in a reasonable amount of computer time. Of course, if we need values for collective quantities such as thermodynamic properties or spatial distribution functions, then the full molecular dynamics method must be used. General discussions of these methods and sample results are provided in the review by Evans [39].

1.7 RIGHT VERSUS WRONG

In previous sections we have discussed how simulation compares with theory and experiment, how computer simulation compares with modeling, and how molecular dynamics compares with other forms of molecular-scale simulation. At this point it would be natural to give sample applications that show the kinds of problems to which molecular dynamics has been applied. However, this last section is not going to give such examples; if you need sample applications, the Bibliography should help you access the literature. Instead of trying to convince you that simulation can produce important results, this section tries to convince you that there is importance in the *way* in which simulation produces results. Those ways—those methods—are in fact worthy of our attention.

1.7.1 Realism, Accuracy, and Validity

To have a context for discussion, we review the reductionistic approach to problems. In reductionism, to predict the behavior of a physical system, we first create a mathematical model. The model mimics the real system only to the extent that it incorporates those features that determine the behavior under investigation. Thus a difficult problem, the real system, is replaced by a simpler problem, the model. We use the term *realism* to mean the extent to which the model reproduces behavior of the real system.

Now the behavior of the model might be calculated via analytic theory. Traditionally, models were kept simple so that their theoretical descriptions could be posed in a closed mathematical form and solved exactly. Models without this simplicity were avoided because, in the absence of computing machines, numerical solutions were too tedious and prone to error. Simple models can usually be accurately described by theory, so comparing a theoretical result with a measurement on the real system provides an unambiguous test of the realism of the model. As suggested in Figure 1.2, such tests advance science through the classic interplay of theory and experiment.

But with the availability of reliable high-speed computers, mathematical models need not be as simple as before and their theoretical descriptions need not be restricted to the analytically solvable. Theories may themselves involve unknown quantities or functions that must be estimated before the theoretical calculation can proceed. Such theories are not necessarily faithful representations of the model; we use the term *accuracy* to refer to the degree to which a theory reproduces behavior of a model. To fix these ideas, the relation $PV = NkT$ is an accurate theoretical result, derivable in statistical mechanics, for the ideal-gas model of matter; however, the ideal gas is an unrealistic model for liquid water.

This new complexity in models and theories means that realism of a model is now an issue separate from accuracy of a theory. Both issues cannot be simultaneously resolved by simply comparing theory with experiment. If the comparison is unfavorable, it may be because the model is unrealistic, or because the theory is inaccurate, or both. If the comparison is favorable, it may be because unrealistic features of the model fortuitously compensate for inaccuracies in the theory.

For determining the behavior of a proposed model, an alternative to analytic theory is provided by computer simulation. However, simulations themselves are not necessarily reliable because they can incur many kinds of errors: uncertainties in reaching an equilibrium or steady state, statistical errors in drawing a finite number of partially correlated samples, numerical errors in the algorithms used to perform the simulation, and round-off errors due to the finite word length of the computer. If, in spite of these errors, the simulated behavior of the model is within some allowable tolerance of its actual behavior, then we say the simulation is *valid*. Therefore, the reductionistic program in Figure 1.1 for studying a real system must confront either the combined effects of realism and accuracy or those of realism and validity; see Figure 1.6.

1.7.2 On the Importance of Being Valid

For the remainder of this section and in the rest of the book, we focus on establishing the validity of a simulation. In some instances, validity can be determined by comparing simulated behavior with that calculated from an analytic theory of known accuracy. Such a comparison constitutes the *strong test* for validity. In exceptional cases we may be able to simulate a model whose behavior is calculable from a completely accurate theory; such cases, in which we demonstrate perfect validity, are called *ideal simulations* [40]. An example is the one-dimensional harmonic oscillator (ODHO). A numerical simulation of the ODHO can be shown to be valid by comparing the simulated motion with that provided by exact analytic theory. However, note that a simulation of the ODHO may not be valid; for example, we might use a crude finite-difference method that erroneously solves the equations of motion. As a model the ODHO may be a realistic representation of the motion of a mass on a spring but be less realistic as a model for the vibratory motion of a carbon bond on a polymer chain.

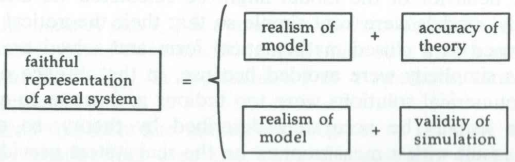

FIGURE 1.6 In a reductionistic study of a real system, we must confront either the combined effects of realism of a model plus accuracy of a theory or of realism of a model plus validity of a simulation.

In practice the strong test for validity is of limited use because we rarely have an analytic theory of known accuracy with which to compare. In fact we often want to use simulation results to test a theory's accuracy rather than use a theory to confirm a simulation's validity. A weaker test for validity is to estimate each of the possible simulation errors and argue that their accumulated effect is within the allowed tolerance for a valid simulation. It is this weak test for validity that must usually be employed and that requires so much attention from the simulator.

Although the validity of a simulation is discussed above in the context of a reductionistic approach to science, note that in fact validity is an issue no matter how the simulation is being used. We must always worry about the reliability of any simulation. And although it is argued in Section 1.3 that simulations are not experiments, nevertheless, the issue of validity is analogous to the issue of accuracy in laboratory measurements. We can probably do no better than take the same attitude to simulation results that the careful experimentalist takes to measurements: What are the possible sources of error or uncertainty? Are sampled populations representative? Are successive samples correlated? Are the results reproducible? Can I do a different simulation that should produce the same results? Are the results internally consistent? Do the results confirm results from previous simulations done elsewhere? How accurate and how precise are the results? In short, simulations done in a deterministic manner on reliable machines are not necessarily correct.

To emphasize this last statement, consider the data in Figure 1.7. Figure 1.7(a) shows simulation results for an observable y obtained at particular values of input observable x. These y-values happen to be from 1983 Monte Carlo runs [41], but neither the kind of simulation nor the identities of x and y affect this lesson. It is sufficient to say that the quantity y is notoriously difficult to extract reliably from molecular simulation and that y should pass through zero when x does. For our purpose the question is, what is the nature of the curve that we would draw through the data in Figure 1.7(a)?

Before proceeding, perform this little experiment: use a sheet of paper (or your hand) to cover panels 1.7(b) and (c) and also cover the rightmost two points in Figure 1.7(a). What kind of curve would you draw through the

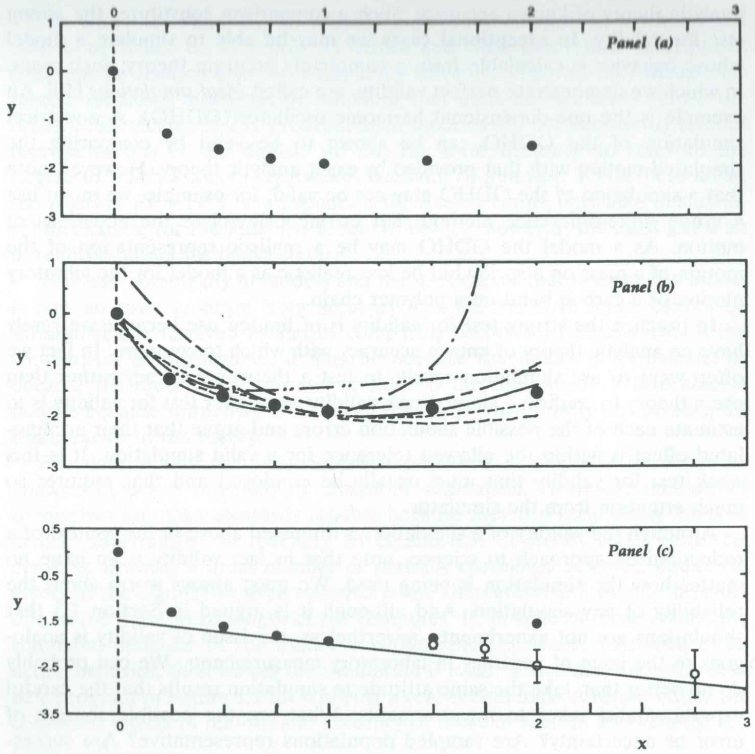

FIGURE 1.7 A single spurious value from simulation can have serious ramifications. *Panel* (*a*): Original Monte Carlo results [41] suggesting a minimum in the quantity $y(x)$. *Panel* (*b*): Results from six efforts to model the simulation data; all show a minimum: —— [42]; ——— [43]; ---- [44]; —··—·· [45]; ---- [46]; —·—· [47]. *Panel* (*c*): Original and additional Monte Carlo results [48] finding that, in fact, $y(x)$ has no minimum. The solid line is from a simple classical theory. The error bars are those provided by the original authors.

remaining five points, a monotone or one with a minimum? I think your answer must be a monotone. Now, shift your paper to the right to expose the next point, the one at $x = 1.5$. Does this additional point change the shape of your imagined curve? Remember, there are uncertainties in these y-values, although the authors haven't given us direct guidance as to the size of the uncertainties. I think your response should be that the point at $x = 1.5$ changes your curve a bit, but its shape is still monotonic. Now uncover the point at $x = 2$. Does this point affect the shape of your imagined curve? If we accept that the y-value at $x = 2$ is as reliable as any of the others (and the authors give us no reason not to), then you conclude that the curve is now changed: the points apparently define a minimum. Note, however, that this conclusion is weakly based because the shape of the curve has been determined primarily by one point, that at $x = 2$. End of experiment. Are you comfortable with a conclusion based on essentially one sample?

The minimum in the data in Figure 1.7(*a*) posed a challenge to modelers. Figure 1.7(*b*) shows various results for y obtained from analytic calculations using several very different models: local composition [42], empirical quartic *PVT* equation of state [43], a kind of mean density approximation [44], variational theory [45], corresponding states [46], and another empirical equation of state [47]. All of these, as well as other models, manage, at least, to generate a minimum.

Now comes the twist in the plot: in 1988 there appeared additional Monte Carlo results for y [48], published by the authors of the 1983 data of Figure 1.7(*a*). These new data are shown, along with the old data, in Figure 1.7(*c*). The new data indicate that the original y-value at $x = 2$ was in error and that y is apparently monotone in x: there is no minimum.

So, what are we to make of the models represented in Figure 1.7(*b*)? Did they really establish a connection between the output y and the input x? Could the authors of those models now adjust parameters so their models fit the data in Figure 1.7(*c*) as well as (or better than) they fit the original data? And if they could make those adjustments, what would they teach us? Would we believe that the revised models now establish connections between y and x? Or are the models merely elaborate curve fits, flexible enough to accommodate many different functional forms?

The lesson here is not that you can necessarily avoid mistakes—mistakes are, after all, a powerful way to learn—rather, the lesson is that you should tread warily. Combining models with simulation can be a tricky and subtle business. To accept simulation results as valid demands careful consideration of the available evidence and often a search for more evidence. Simulation involves more than obtaining a program, performing some runs, and publishing the results. In short, simulations done in a deterministic manner on reliable machines can be wrong.

REFERENCES

[1] W. H. Miller, Ed., *Dynamics of Molecular Collisions*, Parts A and B, Plenum, New York, 1976.

[2] M. Evans, G. J. Evans, W. T. Coffey, and P. Grigolini, *Molecular Dynamics and Theory of Broad Band Spectroscopy*, Wiley, New York, 1982.

[3] G. K. Horton and A. A. Maradudin, Eds., *Dynamical Properties of Solids*, 3 vols., North-Holland, Amsterdam, 1974.

[4] U. Burkert and N. L. Allinger, *Molecular Mechanics*, ACS Monograph 177, American Chemical Society, Washington, DC, 1982.

[5] I. B. Cohen, *Revolution in Science*, Harvard University Press, Cambridge, MA, 1985, Chapter 10.

[6] P. S. Laplace, *Philosophical Essay on Probabilities* (1814); English translation by F. W. Prescott and F. L. Emory (1902), reprinted in *The Beginnings of Modern Science*, H. Boynton, Ed., Walter J. Black, Roslyn, NY, 1948.

[7] W. Thomson (Lord Kelvin), *Baltimore Lectures on Molecular Dynamics and the Wave Theory of Light*, C. J. Clay and Sons (Cambridge University Press), London, 1904. Reprinted as *Kelvin's Baltimore Lectures and Modern Theoretical Physics*, R. Kargon and P. Achinstein, Eds., MIT Press, Cambridge, MA, 1987.

[8] J. L. Casti, *Alternate Realities*, Wiley-Interscience, New York, 1989.

[9] R. Penrose, *The Emperor's New Mind*, Oxford University Press, New York, 1989, Chapter 5.

[10] S. J. Gould, *Wonderful Life: The Burgess Shale and the Nature of History*, Norton, New York, 1989, Chapter 1 and pp. 282–283.

[11] R. F. Fox, *Energy and the Evolution of Life*, Freeman, New York, 1988.

[12] L. Thorndike, *A History of Magic and Experimental Science*, Vol. II, Macmillan, New York, 1929, p. 29.

[13] J. Thurber, "Here Lies Miss Groby," *The Thurber Carnival*, Harper & Brothers, New York, 1945.

[14] A. Funkenstein, *Theology and the Scientific Imagination from the Middle Ages to the Seventeenth Century*, Princeton University Press, Princeton, NJ, 1986, p. 156.

[15] C. G. Gray and K. E. Gubbins, *Theory of Molecular Fluids*, Vol. 1, Oxford University Press, Oxford, 1984.

[16] D. Ceperley and J. Tully, Eds., "Stochastic Molecular Dynamics," *Proc. Nat. Resource Comp. Chem.*, **6**, xiii (1979).

[17] N. A. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of State Calculations by Fast Computing Machines," *J. Chem. Phys.*, **21**, 1087–1092 (1953).

[18] C. Pangali, M. Rao, and B. J. Berne, "On a Novel Monte Carlo Scheme for Simulating Water and Aqueous Solutions," *Chem. Phys. Lett.* **55**, 413–417 (1978).

[19] D. L. Beveridge, M. Mezei, P. K. Mehrotra, F. T. Marchese, G. Ravi-Shanker, T. Vasu, and S. Swaminathan, "Monte Carlo Computer Simulation Studies of the Equilibrium Properties and Structure of Liquid Water," *ACS Adv. Chem. Ser.*, **204**, 297–352 (1983).

[20] M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods*, Vol. 1: *Basics*, Wiley, New York, 1986.

[21] W. W. Wood, "Monte Carlo Studies of Simple Liquid Models" in *Physics of Simple Liquids*, H. N. V. Temperley, J. S. Rowlinson, and G. S. Rushbrooke, Eds., North-Holland, Amsterdam, 1968, Chapter 5.

[22] J. P. Valleau and S. G. Whittington, "A Guide to Monte Carlo for Statistical Mechanics: 1. Highways" in *Statistical Mechanics*, Part A, B. J. Berne, Ed., Plenum, New York, 1977, Chapter 4; J. P. Valleau and G. M. Torrie, "A Guide to Monte Carlo Methods for Statistical Mechanics: 2. Byways," ibid., Chapter 5.

[23] K. Binder, Ed., *Monte Carlo Methods in Statistical Physics*, 2nd ed., Springer-Verlag, Berlin, 1986.

[24] K. Binder, Ed., *Applications of the Monte Carlo Method in Statistical Physics*, Springer-Verlag, Berlin, 1984.

[25] B. J. Alder and T. E. Wainwright, "Studies in Molecular Dynamics. I. General Method," *J. Chem. Phys.* **31**, 459–466 (1959).

[26] J. Ford, "The Transition from Analytic Dynamics to Statistical Mechanics," *Adv. Chem. Phys.*, **24**, 155–185 (1973).

[27] W. C. Swope and H. C. Andersen, "10^6-Particle Molecular-Dynamics Study of Homogeneous Nucleation of Crystals in a Supercooled Atomic Liquid," *Phys. Rev. B*, **41**, 7042 (1990).

[28] A. W. Lees and S. F. Edwards, "The Computer Study of Transport Processes under Extreme Conditions," *J. Phys. C*, **5**, 1921–1929 (1972).

[29] E. M. Gosling, I. R. McDonald, and K. Singer, "On the Calculation by Molecular Dynamics of the Shear Viscosity of a Simple Fluid," *Mol. Phys.*, **26**, 1475–1484 (1973).

[30] W. T. Ashurst and W. G. Hoover, "Argon Shear Viscosity via a Lennard-Jones Potential with Equilibrium and Nonequilibrium Molecular Dynamics," *Phys. Rev. Lett.*, **31**, 206–208 (1973).

[31] W. G. Hoover, "Nonequilibrium Molecular Dynamics," *Ann. Rev. Phys. Chem.*, **34**, 103–127 (1983).

[32] F. Vesely, *Computerexperimente an Flüssigkeitsmodellen*, Physik Verlag, Weinheim, 1978.

[33] W. G. Hoover, *Molecular Dynamics*, Lecture Notes in Physics, **258**, Springer-Verlag, Berlin, 1986.

[34] W. G. Hoover, *Computational Statistical Mechanics*, Elsevier, Amsterdam, 1991.

[35] D. J. Evans and G. Morriss, *Statistical Mechanics of Nonequilibrium Liquids*, Academic, New York, 1990.

[36] R. M. Hockney and J. W. Eastwood, *Computer Simulation Using Particles*, McGraw-Hill, New York, 1981.

[37] G. Ciccotti and W. G. Hoover, Eds., *Molecular-Dynamics Simulation of Statistical–Mechanical Systems*, North-Holland, Amsterdam, 1986.

[38] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Clarendon, Oxford, 1987.

[39] G. T. Evans, "Liquid State Dynamics of Alkane Chains," *ACS Adv. Chem. Ser.*, **204**, 423 (1983).

[40] G. A. Bekey, "Simulability of Dynamical Systems," *Trans. Soc. Comp. Simul.*, **2**, 57–71 (1985).

[41] K. S. Shing and K. E. Gubbins, "The Chemical Potential in Non-Ideal Mixtures. Computer Simulation and Theory," *Mol. Phys.*, **49**, 1121 (1983).

[42] Y. Hu, D. Lüdecke, and J. Prausnitz, "Molecular Thermodynamics of Fluid

Mixtures Containing Molecules Differing in Size and Potential Energy," *Fluid Phase Equilibria*, **17**, 217 (1984).

[43] W. L. Kubic, Jr., "A Quartic Hard Chain Equation of State for Normal Fluids," *Fluid Phase Equilibria*, **31**, 35 (1986).

[44] M. L. Huber and J. F. Ely, "Improved Conformal Solution Theory for Mixtures with Large Size Ratios," *Fluid Phase Equilibria*, **37**, 105 (1987).

[45] E. Z. Hamad and G. A. Mansoori, "Variational Theory of Mixtures," *Fluid Phase Equilibria*, **37**, 255 (1987).

[46] E. C. Meyer, "A One-Fluid Mixing Rule for Hard Sphere Mixtures," *Fluid Phase Equilibria*, **41**, 19 (1988).

[47] B. Smit and K. R. Cox, "A New Approach for Calculating the Accessible Volume in Equations of State for Mixtures. II. Application to Lennard-Jones Mixtures," *Fluid Phase Equilibria*, **43**, 181 (1988).

[48] K. S. Shing, K. E. Gubbins, and K. Lucas, "Henry Constants in Non-Ideal Fluid Mixtures. Computer Simulation and Theory," *Mol. Phys.*, **65**, 1235 (1988).

EXERCISES

1.1 ☒ In later chapters we will have intercourse with one-dimensional periodic systems; an example is the board game Monopoly. The Monopoly universe is a closed line divided into 40 spaces. Play starts from the space labeled Go and moves are clockwise, with the length of each move determined by the throw of two dice. Moves also result from instructions provided by Community Chest and Chance cards.

(a) The strategy used in a game might be enhanced if we knew the spaces most likely to be visited. Write a program that simulates the game and identifies the 10 spaces most likely visited. One player should be sufficient. The program can be developed in stages: (i) first generate moves only by throwing dice, then (ii) add moves prompted by the Community Chest cards, then (iii) add moves prompted by the Chance cards. The random-number generator in Appendix H can be used to simulate throwing dice and shuffling cards.

(b) Does your program simulate Monopoly or does it simulate a model of Monopoly? Justify your answer.

(c) Repeat the calculation in (a), but use only one die; then repeat using three dice.

(d) To determine whether your answer in (a) is affected by initial conditions, repeat the calculation using various starting points other than the Go space.

(e) To determine whether your answer in (a) is invariant under time reversal, repeat the calculation performing moves counterclockwise rather than clockwise.

Of the 40 spaces, the following are either destinations or points of departure for special moves:

Number	Label	Number	Label
2	Community Chest	24	Illinois Ave.
5	Reading RR	25	B&O RR
7	Chance	28	Water Works
10	Jail	30	Go to Jail
11	St. Charles Place	33	Community Chest
12	Electric Co.	35	Short Line RR
15	Pennsylvania RR	36	Chance
17	Community Chest	39	Boardwalk
22	Chance	40	Go

There are 16 Community Chest cards and 16 Chance cards.
Community Chest: Two of these 16 cards direct moves

One Go to Jail (space 10)
One Advance to Go (40)

Chance: Ten of these 16 cards direct moves

One Go to Reading RR (space 5)
One Go to Jail (10)
One Go to St. Charles Place (11)
One Go to Illinois Ave. (24)
One Go to Boardwalk (39)
One Advance to Go (40)
One Go back three spaces
One Advance to nearest utility (12 or 28)
Two Advance to nearest RR (5 or 10 or 15 or 20)

In the last two, note that the instruction is *Advance to nearest* (i.e., move ahead to), not merely Go to.

1.2 The objective of modeling is to decouple interactions; therefore, part of the success of a model hinges on the choice of variables for representing state space, so the desired decoupling is realized. As an example, consider the gravitational two-body problem: masses m_1 and m_2 attract one another with the force model

$$F = -\frac{gm_1m_2}{r^2}$$