1

Introduction – Studying Systems from Two Viewpoints

When analyzing physical, chemical, and biological systems, macroscopic and microscopic viewpoints give two seemingly different descriptions. For example, as shown in Figure 1.1, the state of a gas inside a pressurized capsule is described from a macroscopic viewpoint by a limited number of variables, including pressure, P, volume, V, and temperature, T. Depending on the nature of the gas and the macroscopic state, these variables are related by an equation of state such as the ideal gas law, $PV = nRT$. This description of the gas includes the mechanical variables pressure and volume, and a nonmechanical variable, temperature. In the macroscopic view, the behavior of the gas is governed by the laws of thermodynamics and no reference is made to the molecular structure of the gas. Indeed, thermodynamics was developed in the mid-nineteenth century before atomic theory of matter was widely accepted by physicists. The thermodynamic description is used for macroscopic samples with micrometer or larger length scales over long times.

From a microscopic (atomic) viewpoint, a gas is a collection of a large number (in the order of 10^{23}) of molecules, moving randomly at high speeds, each with a specified position, velocity, and acceleration at any given time. The microscopic description of the gas uses only mechanical variables that obey the laws of classical mechanics. The details of atomic/molecular structures and interactions, along with the application of Newton's equation of motion, determine how positions and velocities of the molecules of the gas change with time. The laws of conservation of energy, linear momentum, and angular momentum constrain the mechanical variables throughout the process. Knowledge of mechanical variables at any time allows the calculation of these variables at all times in the future and past (neglecting considerations of classical nonlinear systems and quantum mechanics) and classical mechanics is therefore *deterministic* with regard to mechanical variables. The classical mechanical microscopic description does not include macroscopic variables such as temperature and entropy, which are used to describe macroscopic systems. The microscopic description is used to describe phenomena on length scales of the order of nanometers and time scales of the order of nanoseconds.

How are these dual descriptions of physical systems reconciled, and why is there such a discrepancy in the length and time scales between these two viewpoints? How do nonmechanical variables get introduced into analysis of

Macroscopic viewpoint
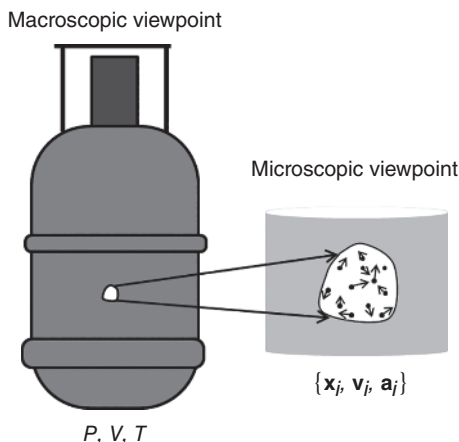
Figure 1.1 Macroscopic and microscopic viewpoints of a gas system involve different variables, length, and time scales.

Microscopic viewpoint

$\{\mathbf{x}_i, \mathbf{v}_i, \mathbf{a}_i\}$

P, V, T

system properties in the macroscopic viewpoint, if these macroscopic variables do not appear in the underlying microscopic description of the system, which is supposedly more fundamental? The answers to these questions form the context of molecular simulation.

In his book "What is Life?" Erwin Schrödinger asks the question every student has wondered when first introduced to atoms: Why are atoms so small? [269] Our daily experience captures length scales as small as millimeters, while atoms and molecules with dimensions in the nanometer range are smaller by factors of $10^{-7}/10^{-8}$ than any phenomena we experience directly. Even the smallest bacteria have dimensions in the micrometer range, which gives them a length range larger by a factor of 10^4 compared to atoms and molecules. Why are there such discrepancies in length and time scales between atoms and the macroscopic phenomenon of life?

Schrödinger argues that since atoms are fundamental building blocks of matter, this is not the correct question to ask. The question should be reframed as "Why are we, as living organisms, so much larger than atoms?" or "Why are there so many atoms and molecules in cells and more complex organisms?" Stated differently, the question can be "Why is Avogadro's number so large?" The answers to these questions determine how new system properties emerge as we transition from the microscopic mechanical descriptions of systems to the macroscopic thermodynamic description of large systems.

The connection between microscopic and macroscopic descriptions is made by invoking probability theory arguments in statistical mechanics. Relatively simple microscopic systems such as ideal gases are amenable to analytical statistical mechanical analysis, and explicit formulas relating microscopic mechanical properties of the gas molecules to macroscopic thermodynamic variables can be derived. For more complex microscopic systems, molecular simulations (using numerical computations) within the framework of molecular dynamics or Monte Carlo simulations are performed and statistical mechanical relations relate the averages of molecular properties to macroscopic observables of these systems.

This book gives an introduction to the microscopic molecular dynamics and Monte Carlo simulation methods for calculating the macroscopic properties of systems. Even in cases where the goal is a purely microscopic mechanical study

of the system, there are usually macroscopic constraints imposed on the system by the environment. For example, the conditions of constant physiological temperature and ambient pressure imposes constraints on molecular simulations when studying the interaction of a drug candidate with an enzyme binding site in aqueous solution. These constraints impose nonmechanical conditions on the microscopic description of the system that must be applied correctly when simulating molecular behavior.

Chapter 2 gives a brief overview of classical mechanics used to describe the motion of atoms and molecules in microscopic systems. We start from simple physical systems for which analytical solutions of the classical Newtonian equations are available and move to complex multiatom systems for which numerical methods of solution (namely, finite difference methods) are needed. The concept of phase space trajectory, which describes the dynamics of these systems, is introduced.

Solving Newton's laws of motion for a molecular system requires knowledge of the forces acting between atoms. In Chapter 3, the quantum mechanical basis for determining the interatomic forces within and between molecules and their classical approximations are described. A description of classical force fields used in molecular simulations of chemical and biological systems follows.

Having introduced numerical methods to solve the classical equations of motion and the microscopic forces acting between atoms, the next step is the introduction of specialized techniques needed to make molecular simulations feasible. These techniques, which include the use of periodic boundary conditions, potential cutoffs for short range forces, and Ewald summation methods for long range electrostatic forces, are discussed in Chapter 4.

In Chapters 5 and 6, we introduce concepts from probability theory that describe how to predict and analyze behaviors of complex systems on which we have too little or too much information. Concepts of probability theory as applied to mechanical systems form the framework for statistical mechanics. Relations of probability theory and statistical mechanics must be considered to correctly run a molecular simulation and to ensure that the molecular level system is treated in a manner consistent with macroscopic conditions imposed on the system. Molecular simulation results can then be subjected to further statistical mechanical analysis or be used to get direct microscopic insight into phenomena of interest. In Chapter 5, the principles of probability theory are applied to non-interacting systems, while in Chapter 6, the concept of the ensemble of systems is introduced, which allows probabilistic analysis of systems that include intermolecular interactions. The classical expressions for the probability distributions for different ensembles are the constraints that molecular simulations must satisfy.

Chapters 7 and 10 cover specialized molecular simulation techniques for imposing specific values of macroscopic thermodynamic variables in a simulated system. In Chapter 7 methods of correctly imposing constant pressure (Andersen barostat) and constant temperature (Nosé–Hoover thermostat) on systems of molecules in molecular simulations are described. In Chapter 10, the grand canonical Monte Carlo simulation method for imposing the condition of fixed chemical potential and temperature is described.

Chapter 8 and 9 treat the extraction and analysis of structural/thermodynamic properties and dynamic properties using molecular dynamics simulations, respectively. Selected examples from a large body of simulation work are outlined.

Throughout the book, we will emphasize an appreciation of time, length, and energy scales of molecular processes including molecular translations, vibrational processes, and bulk fluid motions.

Many excellent books, articles, and websites on mechanics, probability theory, statistical mechanics, and molecular simulation methods are available and have been cited in the references. These have undoubtedly influenced the presentation of the material here and explicit citations are given in different sections as appropriate. A large body of work on molecular dynamics and Monte Carlo methods is available and only a small sample of topics could be covered here. Important and groundbreaking work by many experts has not been discussed, and this is a reflection of the limited scope of this book rather than the importance of the work. Contributions of researchers from the past and present are gratefully acknowledged, although they are not mentioned individually here.

A further point is that many important advanced modern topics are not covered in this book as they are beyond the scope of this introductory discourse. For example, free energy methods, biased Monte Carlo sampling, and methods of high-performance computing used in molecular simulations are not discussed. It is hoped that the introductory material in this book provides a launching pad for the study of these advanced topics. For more advanced users, it is hoped that this book can provide a useful overview and some intuitive understanding of methods that go into molecular simulations.