1

Microscopic Origin of Thermodynamics

The real-life applications of thermodynamics hinge on experimental measurements and/or theoretical predictions of thermodynamic properties for macroscopic systems of practical interest. The experimental approach is exemplified by the traditional applications of engineering and chemical thermodynamics as evidenced in the extensive use of thermodynamic tables, diagrams, and semi-empirical correlations. Conversely, the theoretical approach is based on statistical mechanics, which is of central concern in this book.

This introductory chapter presents the key hypotheses of statistical mechanics to describe the thermodynamic properties of an equilibrium system from a microscopic perspective. After a brief overview recapitulating the essential ingredients of classical thermodynamics and fundamental relations linking different thermodynamic variables, we introduce *internal energy* and *entropy* – two most fundamental quantities of thermodynamics – in terms of the properties of individual particles, i.e., the *microscopic constituents* of a thermodynamic system. The *statistical nature* of thermodynamic variables will be elucidated in the context of ensemble, ergodicity, and microstates. In addition, we will discuss *flexibility* in the statistical descriptions of individual particles and microstates of a thermodynamic system, as well as *additivity* and *relativity* pertaining to both internal energy and entropy when assessed from microscopic perspectives.

We assume that the readers are already familiar with the fundamentals of classical thermodynamics including their applications to chemical and phase equilibria. However, the previous exposure of statistical mechanics is not expected. The supplementary material gives a brief overview of classical and quantum mechanics for those who are unfamiliar with this subject. While an advanced understanding of quantum mechanics is not a prerequisite, basic concepts such as Hamiltonian, quantum states, and the Schrödinger equation will be used to describe particle energy and the microscopic constituents of quantum systems (e.g., photons, electrons, and phonons).

1.1 Microscopic Constituents of Thermodynamic Systems

In this section, we discuss the essential ideas of classical thermodynamics from a microscopic perspective and recapitulate the fundamental relationships among different thermodynamic variables such as temperature, pressure, entropy, and energy. In addition, we will elucidate how statistical mechanics helps to understand the macroscopic properties of thermodynamic systems based on the dynamic behavior of their constituent particles.

Fundamentals and Practice in Statistical Thermodynamics, First Edition. Jianzhong Wu and John M. Prausnitz.

Companion website: www.wiley.com/go/jianzhongwu/fundamentalsandpractice

1.1.1 Classical Thermodynamics

Classical thermodynamics is centered around two fundamental laws of nature that are universally applicable to the *collective* behavior of macroscopic systems. The first law follows the conservation of the total energy, and the second law asserts that spontaneous events in nature proceed in a particular direction. These thermodynamic laws were established in the second half of the nineteenth century from repeated observations of natural phenomena underlying transformations among different forms of energy and their connections with the physical properties of matter. Extensive experience over many years renders us confidence that the thermodynamics laws are permanent, unlikely to be refuted by future scientific developments. As famously stated by Albert Einstein,[1] "*A theory is the more impressive the greater the simplicity of its premises is, the more different kinds of things it relates, and the more extended is its area of applicability. Therefore the deep impression which classical thermodynamics made upon me. It is the only physical theory of universal content concerning which I am convinced that, within the framework of the applicability of its basic concepts, it will never be overthrown (for the special attention of those who are skeptics on principle).*"

Closely affiliated with the fundamental laws of thermodynamics are two indispensable quantities, internal energy U and entropy S. As discussed in more detail later in this chapter, both internal energy and entropy are defined by the *microscopic constituents* of a thermodynamic system. More precisely, internal energy refers to the total energy arising from the perpetual motions of *individual particles* and inter-particle interactions. The former is commonly known as *kinetic energy*, and the latter is *potential energy*. Conversely, entropy provides a measure of the uncertainty of a macroscopic system in terms of the dynamic behavior of the individual particles or, without concerning the time, in terms of *possible ways* that individual particles may exist (e.g., particle positions and momenta or wave functions) under a particular thermodynamic condition. Establishing the connection between thermodynamic quantities and the dynamic behavior of individual particles is an essential task of statistical thermodynamics.

Classical thermodynamics is concerned with variations in the *equilibrium* properties of macroscopic systems, i.e., systems consisting of many particles, typically on the order of 10^{23}. For a macroscopic system at equilibrium, all thermodynamic quantities are fixed.[2] In other words, all macroscopic properties of interest are independent of time. Here, time independence means that the duration of observation is sufficiently long compared with the time scale that characterizes the dynamics of individual particles. Thermodynamic laws cannot be applied if one is interested only in a single particle in the vacuum or even a few particles. If a system contains only a few particles, the dynamic behavior can be described with conventional equations from classical or quantum mechanics. It is the enormous number of particles in a thermodynamic system that prevents the direct use of the mechanical equations to describe particle motions at a level of certainty the same as that for a system with only a small number of particles. *Uncertainty at the microscopic level* is intrinsic for all thermodynamic systems.

The individual particles of a thermodynamic system embody not only kinetic and potential energies but also information concerning the precise meaning of microscopic constituents. Thermodynamics makes no assumption on the physical nature of the individual particles, i.e., the thermodynamic laws hold regardless of how the microscopic constituents are characterized or interpreted. Indeed, individual particles in a thermodynamic system are rather diverse; they may

---

[1] Einstein A., "Autobiographical Notes", p. 33, in *Albert Einstein: Philosopher-Scientist*, Schilpp P. A., Ed., The Library of Living Philosophers Volume VII. MJF Book, New York, 1970.

[2] The macroscopic properties of a steady-state system are also independent of time, but that system is not necessarily at equilibrium. In a steady-state system, there is a net flux of energy or mass, independent of time. However, for equilibrium, it is necessary that there is no net flux of mass or energy is zero.

refer to atoms or molecules as commonly present in different states of matter (i.e., gases, liquids, and solids), or elementary particles such as photons and electrons, or certain aspects of elementary particles (e.g., magnetic spins), or individual units of a molecule (segments), or aggregates of molecules (e.g., colloidal particles). The *flexibility* in interpreting the microscopic constituents of a thermodynamic systems suggests that thermodynamic variables such as internal energy and entropy are intrinsically *relative*, i.e., their absolute values are dependent on the *definition* of individual particles. For example, internal energy and entropy may take different values when an atomic system is described in terms of electrons and atomic nuclei with quantum mechanics or as individual atoms like classical particles. In either case, the changes in thermodynamic properties predicted from a microscopic perspective should be in accordance with experimental observations.

1.1.2 The Fundamental Equation of Thermodynamics

In addition to internal energy and entropy, we use *auxiliary* quantities such as enthalpy H, Helmholtz energy F and Gibbs energy G to describe the macroscopic properties of a thermodynamic system. These auxiliary functions are introduced for the convenience of practical applications when thermodynamic systems are prepared under different circumstances (e.g., fixed total energy or fixed temperature, constant pressure or constant volume, etc.). All auxiliary quantities can be *formally* derived from internal energy and entropy along with variables that specify a thermodynamic system (e.g., temperature, pressure, and total volume). Using the auxiliary functions, we can describe heat effects for constant pressure processes simply in terms of the changes in enthalpy, and the thermodynamic limits of various isothermal processes as well as the conditions of equilibrium with different thermodynamic potentials or free energies.

For a closed system, i.e., a system free of mass transfer with its surroundings, classical thermodynamics asserts that the internal energy U and the entropy S are related through the fundamental equation

$$dU = TdS - PdV. \tag{1.1}$$

Eq. (1.1) can be obtained by applying the first and second laws to a closed system undergoing a reversible process that involves volumetric work $-PdV$ and heat transfer TdS. Additional variables must be introduced to account for other forms of reversible work or the mass transfer of any chemical species between the system and its surroundings. According to multivariable calculus, the differential form in Eq. (1.1) suggests that the internal energy U is an analytical function of entropy S and volume V, i.e., $U = U(S, V)$. In addition, Eq. (1.1) implies that temperature T and pressure P can be expressed as partial derivatives

$$T = \left(\frac{\partial U}{\partial S}\right)_V, \tag{1.2}$$

$$P = -\left(\frac{\partial U}{\partial V}\right)_S. \tag{1.3}$$

In practical applications, a closed system is often specified in terms of temperature and volume or temperature and pressure. Accordingly, it is convenient to describe the thermodynamic properties with such variables. Mathematically, the change of independent variables can be achieved through the Legendre transformation[3]:

$$F \equiv U - \left(\frac{\partial U}{\partial S}\right)_V S = U - TS, \tag{1.4}$$

---

[3] The Legendre transformation of an analytical function results in a new function in which the slope of the original function becomes the independent variable. For example, $f(x) \to g(y) \equiv f(x) - yx$, where $y = f'(x)$, represents the Legendre transformation of function $f(x)$ to $g(y)$. The transformation implies $dg = f' dx - d(yx) = -xdy$, i.e., $-x = g'(y)$, where g is a function with y as the independent variable.

$$H \equiv U - \left(\frac{\partial U}{\partial V}\right)_S V = U + PV, \tag{1.5}$$

$$G \equiv H - \left(\frac{\partial H}{\partial S}\right)_P S = H - TS \tag{1.6}$$

where $F = F(T, V)$ is called Helmholtz energy, $H = H(S, P)$ is enthalpy, and $G = G(T, P)$ is Gibbs energy. Substituting Eqs. (1.4)–(1.6) to (1.1) leads to alternative forms of the fundamental equation

$$dF = -SdT - PdV, \tag{1.7}$$

$$dH = TdS + VdP, \tag{1.8}$$

$$dG = -SdT + VdP. \tag{1.9}$$

The procedure can be generalized to other forms of thermodynamic potentials (e.g., grand potential to be discussed in Chapter 2), and similar fundamental equations can be derived by Legendre transformation.

While the formal relations discussed above are most relevant to gases or liquids in a bulk phase, similar equations can be readily established for two- and one-dimensional systems such as molecules adsorbed at a surface or in narrow micropores. For a two-dimensional system, the thermodynamic properties depend on surface area A instead of volume V. Accordingly, the fundamental equation is given by

$$dU = TdS - \varsigma dA \tag{1.10}$$

where ς denotes the surface pressure, i.e., the variation of energy with the surface area. For a one-dimensional system, the system length is often used as the spatial variable. For example, the fundamental equation for a rubber band can be expressed as

$$dU = TdS + fdL \tag{1.11}$$

where f and L represent the tension and length of the rubber band, respectively. Similar to those for bulk fluids, other forms of the fundamental equation can be derived for both two- and one-dimensional systems.

1.1.3 Statistical Thermodynamics

Statistical mechanics aims to predict the properties of a macroscopic system based on the dynamic behavior of individual particles. To establish connections between micro- and macroscopic properties of a thermodynamic system, statistical mechanics offers rigorous mathematical procedures to evaluate the *statistical* distributions of the microstates of individual particles. Such procedures are universally applicable to all many-body systems, i.e., from systems containing a few particles up to the thermodynamic limit.[4]

In statistical mechanics, a *microstate* refers to a particular way that the individual particles of a thermodynamic system exist. In other words, at any microstate, all variables concerning the dynamic behavior of individual particles are fully specified. If the particles are classical and have a spherical shape, these variables correspond to the positions and momenta of individual spheres. If the particle motions are described by quantum mechanics in terms of the wave function of the entire system, each microstate may be understood as a quantum state. As mentioned above, individual particles in a thermodynamic system are not necessarily correspondent to the

---

[4] In the thermodynamic limit, the system size, including the number of particles, approaches infinity.

fundamental constituents of matter as defined in particle physics. In introducing microstates, we assume that the dimensionality of individual particles is exceedingly small in comparison to that of a macroscopic system. At any moment, a thermodynamic system may reside in one of many microstates.

According to statistical mechanics, the mechanical behavior of individual particles is responsible for all macroscopic properties of a thermodynamic system, including not only internal energy and entropy but also all auxiliary properties such as chemical potential, enthalpy, and the equilibrium constants of chemical reactions. As uncertainty is inevitable in describing the dynamic behavior of individual particles, statistical mechanics predict the macroscopic properties in terms of the *statistical averages* of quantities related to the mechanical behavior of individual particles. By accounting for the probability of a macroscopic system in different microstates, statistical mechanics can predict, in principle, both static and dynamic properties including those corresponding to systems at nonequilibrium conditions. While the procedure is applicable to macroscopic systems at or away from equilibrium in principle, our concern in this book is focused on the microscopic structure and thermodynamic properties of equilibrium systems.

1.1.4 Summary

In this section, we recapitulate some of the most important concepts from classical thermodynamics, the microscopic meaning of key thermodynamic quantities such as internal energy and entropy, and the connection between thermodynamics and statistical mechanics. While thermodynamics is, as Einstein indicated, a theory of principle that is universally applicable to any macroscopic systems, statistical mechanics intends to establish connections between thermodynamic variables with the microscopic constituents of thermodynamics systems. Statistical thermodynamics is a branch of statistical mechanics that is concerned with the properties of equilibrium systems. In this book, we do not distinguish these two terms and use them interchangeably because conventional statistical-mechanical methods are applicable only at or near equilibrium conditions.

1.2 Thermodynamic Relations

In this section, we analyze the mathematical implications of thermodynamic functions and discuss how they may lead to formal thermodynamic relations such as the Gibbs–Duhem and the Maxwell relations. These relations are useful because they facilitate easier application of thermodynamic principles and enable quantitative predictions of thermodynamic properties that may not have obvious microscopic counterpart or be directly accessible through experiments. Exact relations among different thermodynamic quantities are also useful for testing the self-consistency of experimental and/or theoretical results.

1.2.1 Intensive and Extensive Variables

Some thermodynamic variables, such as temperature, pressure, volume, and composition, can be measured directly; while others, such as entropy, internal energy, and chemical potential, are determined indirectly from measurements of other, directly measurable, variables. Both types of thermodynamic variables can be intensive or extensive. An intensive variable is independent of the system size, and an extensive variable is size-dependent.

Examples of intensive variables are temperature T, pressure P, and the chemical potentials of individual chemical species μ_i. These variables are associated with the driving forces for the exchange of energy and mass between the system and its surroundings: T dictates the direction of heat transfer, P is related to the volumetric work, and μ_i reflects the tendency of mass transfer through diffusion or chemical conversion. Extensive variables include internal energy U, entropy S, enthalpy H, Helmholtz energy F, Gibbs energy G, constant-pressure heat capacity C_P or constant-volume heat capacity C_V, the number of moles n_i or mass m_i for species i, and the system volume V. Other extensive variables are possible when we consider nonvolumetric work such as that due to the change in interfacial area or particle motions in an external field. Extensive variables are linearly scaled with the system size, i.e., if a system is replicated a certain number of times, all extensive properties are multiplied by the same factor accordingly.

The linear dependence of extensive variables on the system size implies that they can be converted into intensive variables after being normalized with the total mass, volume, or the number of moles of all chemical species. For example, we can obtain specific internal energy and specific entropy by dividing the total internal energy and the total entropy by the system mass, respectively. Similarly, molar internal energy and molar entropy correspond to the total internal energy and the total entropy divided by the total number of moles, and the internal energy density and the entropy density correspond to the total internal energy and the total entropy divided by the system volume.

1.2.2 The Gibbs Phase Rule

One remarkable feature of classical thermodynamics is that an equilibrium system can be defined by a few thermodynamic variables regardless of its complexity. For example, except for a scaling factor related to the system size, the properties of a uniform system such as liquid water in the bulk phase[5] are completely specified by temperature and pressure. Similarly, a multi-component system may be specified by temperature, pressure, and composition if it exists in a single phase. In practice, the composition may be defined in terms of mole fractions, volume fractions, weight fractions, or the concentrations of different chemical species. These variables are both necessary and sufficient to characterize the properties of a macroscopic system at equilibrium. Once a system is defined by these few thermodynamic variables, all macroscopic properties of the system can be determined, in principle, by either experiment or theoretical calculation.

For a multicomponent system free of chemical reactions, the Gibbs phase rule predicts that a thermodynamic state can be fully specified by N_f *independent* intensive variables

$$N_f = \mathcal{C} - \mathcal{P} + 2 \tag{1.12}$$

where $\mathcal{C}$ is the number of chemical species or components in the mixture, and $\mathcal{P}$ is the number of coexisting phases. The number of independent variables to define a thermodynamic system is called *the degrees of freedom*. Eq. (1.12) was derived first by J. W. Gibbs by applying the first and second laws of thermodynamics to macroscopic systems at phase and chemical equilibria.[6]

---

[5] In thermodynamics, a phase refers to a state of matter such as gas, liquid or solid that can be distinguished from other states in terms of both the microscopic structure and macroscopic properties.

[6] Josiah Willard Gibbs published the paper entitled "On the Equilibrium of Heterogeneous Substances" in two parts in the *Transactions of the Connecticut Academy of Arts and (3, 108–248, Oct 1875–May 1876 and 343–524, May 1875–1878). Little known for many years after its publication, the 300-page paper is sometimes referred to as "the principia of thermodynamics" (Wikipedia). Morowitz considered it as "the second most significant document produced in the United States" (Morowitz, H. J., "Let free energy ring", Hospital Practice, 11:189–190, 1976).*

It is worth noting that, in a thermodynamic system, intensive variables may be implicitly related to each other. For example, the molecular fractions of a multicomponent mixture, $x_{i=1,2,\ldots C}$, are intensive variables but not fully independent of each other. The mole fractions are normalized, i.e.,

$$\sum_{i=1}^{C} x_i = 1. \tag{1.13}$$

Therefore, the number of independent variables to describe the composition of an C-component system is $C - 1$.

The Gibbs phase rule does not apply to variables with intrinsic cause and effect relations, such as pressure P and volume V, temperature T and entropy S, chemical potential μ_i and the number of molecules N_i for chemical species i. For example, because liquid water exhibits a maximum density near 4 °C and $P = 1$ atm, specifying pressure and molar volume may not be sufficient to define the thermodynamic state. As shown schematically in Figure 1.1, two thermodynamic states I and II may have the same density and pressure. In thermodynamics, two variables with a cause-and-effect relation are known as a conjugate pair. Typically, a conjugate pair consists of an intensive variable and an extensive variable per unit volume, mass, or mole. The former serves as the thermodynamic driving force for the change of an extensive variable (e.g., P, T and μ_i are the driving forces for the changes in V, S, and N_i, respectively).

While the Gibbs phase rule can be similarly formulated for all thermodynamic systems, Eq. (1.12) is valid only for systems containing multiple bulk phases, i.e., for multi-phase systems free of chemical reactions. In that case, the intensive variables for each phase can be completely specified by temperature, pressure, and chemical composition. In general, a thermodynamic system may also be subject to external potential, i.e., energy related to the positions of individual particles. For example, any thermodynamic system on Earth experiences a gravitational potential due to the mass of its particles. When molecules are confined within porous materials, they are subject to interactions with the solid structure. In such circumstances, the properties of a thermodynamic system are defined not only by the characteristics of the chemical species under bulk conditions but also by the parameters that describe the external potential.

1.2.3 The Gibbs–Duhem Equation

The multiplicative behavior of extensive variables implies that thermodynamic quantities must be interrelated. If the size of a thermodynamic system is scaled by a factor λ, a positive number, all extensive properties of the system must be scaled by the same factor. The linear scaling behavior leads to the Gibbs–Duhem equation, which relates the variation of different thermodynamic quantities.

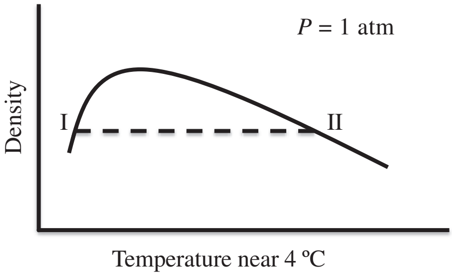


Figure 1.1 Gibbs phase rule predicts the degree of freedom for a one-component system in a single phase is $N_f = 2$. But why are pressure and density insufficient to define liquid water near 4 °C?

To elucidate the mathematical implications of extensive variables, consider a uniform one-component fluid as an example. Similar equations can be readily established for other systems. For a one-component fluid such as a gas or a liquid, the fundamental equation suggests that the internal energy U may be expressed as an analytical function of entropy S, volume V, and the number of molecules N, i.e., $U = U(S,V,N)$. Because these variables are all extensive, their linear scaling by a factor λ leads to

$$\lambda U = U(\lambda S, \lambda V, \lambda N). \tag{1.14}$$

In mathematics, the linear relation among multiplicative variables as shown in Eq. (1.14) is known as a linearly homogeneous function. Accordingly, Euler's theorem predicts

$$U = S\left(\frac{\partial U}{\partial S}\right)_{V,N} + V\left(\frac{\partial U}{\partial V}\right)_{S,N} + N\left(\frac{\partial U}{\partial \mu}\right)_{V,S} = TS - PV + \mu N. \tag{1.15}$$

Eq. (1.15) can be obtained by differentiating both sides of Eq. (1.14) with respect to λ and thermodynamic relations

$$T = \left(\frac{\partial U}{\partial S}\right)_{V,N}, \quad P = -\left(\frac{\partial U}{\partial V}\right)_{S,N}, \quad \text{and} \quad \mu = \left(\frac{\partial U}{\partial N}\right)_{V,S}. \tag{1.16}$$

Recall that the fundamental equation predicts

$$dU = TdS - PdV + \mu dN. \tag{1.17}$$

A combination of Eqs. (1.15) and (1.17) leads to the Gibbs–Duhem equation for a one-component system

$$0 = SdT - VdP + Nd\mu. \tag{1.18}$$

As the Gibbs phase rule predicts that there are only two independent variables for the one-component system, intensive variables T, P, and μ must be related to each other. Specifically, Eq. (1.18) predicts that, at constant temperature, the variation in pressure with respect to chemical potential gives

$$\left(\frac{\partial P}{\partial \mu}\right)_T = \rho \tag{1.19}$$

where $\rho = N/V$ is the number density of molecules. Accordingly, the isothermal compressibility can be written as

$$\kappa_T \equiv \frac{1}{\rho}\left(\frac{\partial \rho}{\partial P}\right)_T = \frac{1}{\rho^2}\left(\frac{\partial \rho}{\partial \mu}\right)_T. \tag{1.20}$$

Eqs. (1.19) and (1.20) allow us to calculate the variation of the chemical potential with respect to pressure or density. Besides, it can also be used to establish, as discussed in later chapters (Sections 2.9 and 7.3), connections between compressibility and fluctuations of thermodynamic quantities.

Other forms of the Gibbs–Duhem equation can be readily derived from Euler's theorem. For example, for a uniform multicomponent system at fixed temperature and pressure, any extensive property can be expressed as a function of the number of moles for each chemical species, $X = X(T, P, n_i)$. The Gibbs–Duhem equation predicts that, at constant T and P,

$$0 = \sum_i n_i d\overline{X}_i \tag{1.21}$$

where $\overline{X}_i = (\partial X/\partial n_i)_{T,P,n_{j\neq i}}$ is called partial molar property. Note that chemical potential for each species is a partial molar property, $\mu_i = \overline{G}_i = (\partial G/\partial n_i)_{T,P,n_{j\neq i}}$. Thus, a special form of the Gibbs–Duhem equation is given by

$$0 = \sum_i n_i d\mu_i. \tag{1.22}$$

Similar equations can be deduced for two-dimensional and one-dimensional systems.

1.2.4 The Maxwell Relations

We can obtain the Maxwell relations by observing that the second-order partial derivatives of a multivariable function are independent of the sequence of differentiation. Using again one-component systems as an example, the second-order partial derivative of $U = U(S,V,N)$ with respect to S and V leads to

$$\left(\frac{\partial T}{\partial V}\right)_{S,N} = -\left(\frac{\partial P}{\partial S}\right)_{V,N}. \tag{1.23}$$

A more useful Maxwell relation can be derived from the fundamental equation

$$dF = -SdT - PdV + \mu dN \tag{1.24}$$

by applying the second-order partial derivative of $F = F(T,V,N)$ with respect to T and V

$$\left(\frac{\partial S}{\partial V}\right)_{T,N} = \left(\frac{\partial P}{\partial T}\right)_{V,N}. \tag{1.25}$$

Other Maxwell relations can be readily derived from Eq. (1.24)

$$\left(\frac{\partial S}{\partial N}\right)_{T,V} = -\left(\frac{\partial \mu}{\partial T}\right)_{V,N}, \tag{1.26}$$

$$\left(\frac{\partial P}{\partial N}\right)_{T,V} = -\left(\frac{\partial \mu}{\partial V}\right)_{T,N}. \tag{1.27}$$

Given an equation of state relating pressure P with temperature T, volume V, and the number of particles N, Eq. (1.25) allows us to predict how the entropy of a closed system, an abstract quantity not directly measurable in experiments, changes with volume at constant temperature.

Many more Maxwell relations can be derived from other forms of the fundamental equation including those for multi-component systems. In practice, the most useful Maxwell relations are those that relate abstract quantities such as entropy and chemical potential to thermodynamic variables that can be measured in the laboratory.

1.2.5 Summary

The Gibbs phase rule predicts the number of independent variables to specify the equilibrium state of a macroscopic system. Because all other variables are dependent on a few variables that define the equilibrium system, thermodynamic quantities are interrelated through a network of equations. While the discussion above is focused on bulk fluids, similar relations can be obtained for other thermodynamic systems including those with different dimensionality. Such relations provide exact connections among different thermodynamic variables. Experimental or theoretical results that violate any exact thermodynamic relation are known as *thermodynamic inconsistency*, which should be avoided in practical applications.

1.3 Microscopic Uncertainty, Ensemble Average, and Ergodicity

In classical thermodynamics, each equilibrium state is defined by a few macroscopic quantities as predicted by the Gibbs phase rule. At any instant, a thermodynamic system may reside in one of its many microstates. Unless there are experimental means to keep track of the dynamic behavior of individual particles, uncertainty is inevitable in describing the microscopic details of a thermodynamic system.

The microscopic uncertainty of thermodynamic systems necessitates that the connection between macroscopic properties and the variables underlying the mechanical properties of individual particles can be established only through *statistical* means. To account for microscopic uncertainty, we introduce in this section ensemble average and ergodicity, two most fundamental concepts in statistical mechanics.

1.3.1 Microscopic Uncertainty

Equilibrium systems are considered thermodynamically equivalent if they exhibit the same macroscopic properties. From a microscopic point of view, however, systems with identical thermodynamic properties can be quite different in terms of the dynamic behavior of individual particles.

To illustrate, Figure 1.2 shows two thermodynamic systems, A and B, containing argon gas with the same temperature, pressure, and total volume. If the containers are identical, we assert that all macroscopic properties of system A are the same as those of B, i.e., any measurement for A would be equivalent to that for B. But at any instant, do they have the same energy? Do individual argon molecules have the same relative positions? Do they have the same momenta? Classical thermodynamics is not concerned with these, and other questions related to the microscopic behavior of individual particles. However, such information is essential in statistical thermodynamics because it aims to predict macroscopic properties based on microscopic details.

One mole of argon gas contains about 6.02×10^{23} molecules. As shown schematically in Figure 1.2, argon molecules are in perpetual motion and interact with each other and with the surface of the containers. At any instant, it is extremely unlikely that all argon molecules in

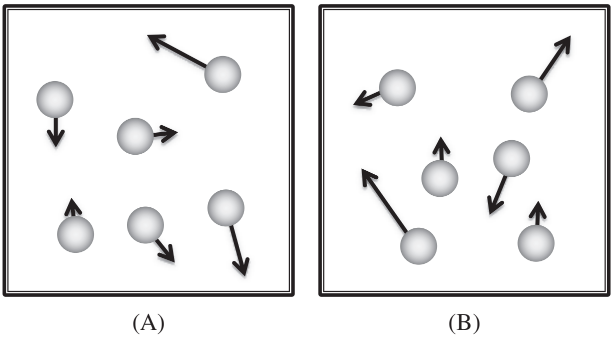


Figure 1.2 Two thermodynamic systems containing argon gas are considered identical if they have the same temperature, pressure, and total volume. But they may be differentiated from a microscopic perspective because, at any instant, the relative positions, and momenta of argon molecules in system (A) are different from those in system (B). The thermodynamic systems of argon gas are equivalent only on a macroscopic scale, not necessarily on the microscopic scale of individual molecules. Schematically, here each sphere represents an argon molecule, and the arrow denotes the direction and magnitude of momentum.

one system have the positions and momenta the same as those in the other even when they are thermodynamically equivalent. In other words, while the thermodynamic (or macroscopic) state of the argon gas in system A is the same as that in system B, at any instant, the microstate in A is most likely different from that in B.

The above example illustrates that, in application of statistical mechanics to any system of practical interest, we need to define the microstates using a physical model suitable for characterizing the microscopic details. In addition, we must determine the probability distribution for the microstates of the system under consideration. With a precise knowledge of the microstates and their probability distribution, we can then predict macroscopic properties by taking statistical averages.

1.3.2 Ensembles

Ensemble plays a vital role in the statistical description of microstates and in establishing their connections with thermodynamic quantities. Formally, a statistical ensemble, or *ensemble* in short, is referred to as an arbitrarily large number of thermodynamic systems that are macroscopically equivalent.

Figure 1.3 shows a schematic picture of an ensemble of thermodynamic systems consisting of identical particles with the same temperature T, volume V, and particle number N. At any instant, each system has a particular microstate, which may or may not be the same as that of other systems in the ensemble. As the number of macroscopically equivalent systems can be arbitrarily large, the ensemble includes not only all microstates of a real system under study, but also encompasses microstates corresponding to any “mental copies” of the real system, i.e., imaginary systems with thermodynamic properties identical to those of the experimental system under investigation.

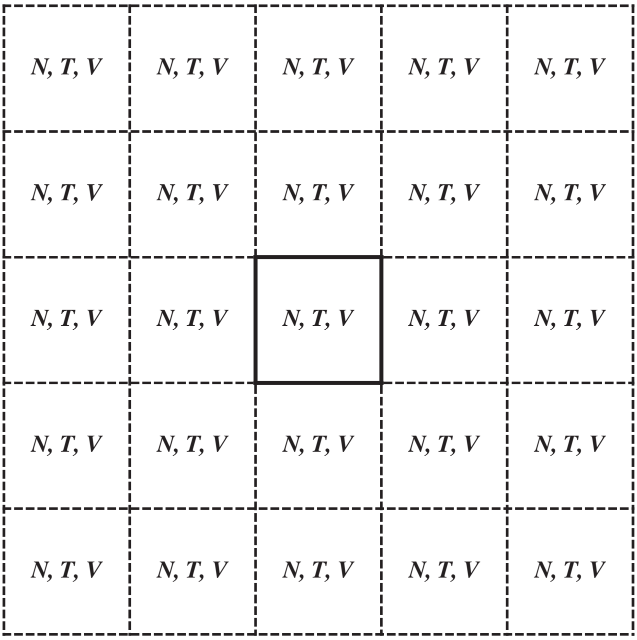


Figure 1.3 An ensemble is defined by an arbitrarily large number of thermodynamic systems with identical macroscopic properties, e.g., temperature T, the number of particles N, and volume V. Here, the box in solid lines represents a thermodynamic system under study; other boxes are mental copies of the real system. All systems have the same T, N, and V. However, at any moment, individual systems may (but need not) exist in microstates different from each other.

By accounting for an arbitrarily large number of systems that are thermodynamically equivalent but not microscopically identical, statistical mechanics provides mathematical means to determine the probability of all microstates in the ensemble and evaluate the average values of both microscopic and macroscopic properties of the real system under study. We will discuss different definitions of microstates in Sections 1.5 and 1.6. Chapter 2 gives formal mathematical procedures to determine the microstate probability distributions and their connections with various thermodynamic quantities based on the dynamic behavior of individual particles. The probability distributions may be alternatively evaluated by counting microstates in many systems consistent with the thermodynamic conditions of interest (e.g., by using a supercomputer), much like polling the approval rate of a politician. This procedure can be implemented with Monte Carlo simulation, which will be discussed in Chapter 6.

1.3.3 Ensemble Averages

An ensemble average refers to the mean value of any quantity affiliated with the microstates of a thermodynamic system. Suppose that M_ν is a property that can be determined for each microstate ν (e.g., the total energy), the ensemble average of M_ν is defined as

$$\langle M \rangle = \sum_\nu p_\nu M_\nu \tag{1.28}$$

where p_ν is the probability of microstate ν in the ensemble, and brackets $\langle \cdots \rangle$ denote an ensemble average.[7] Like any probability distribution, the microstate probability is nonnegative and normalized, i.e.,

$$\sum_\nu p_\nu = 1. \tag{1.29}$$

The ensemble average is applicable to either microscopic or macroscopic quantities, i.e., to the properties of individual particles or of the entire system. For example, internal energy is defined as the ensemble average of the total kinetic and potential energies of all particles over all microstates

$$U = \sum_\nu p_\nu E_\nu \tag{1.30}$$

where E_ν is the total energy of the system at microstate ν. It should be noted that E_ν is associated with the energies of the microscopic constituents (viz., particles) of a thermodynamic system; it does not include the kinetic energies due to the overall motion or potential energies related to the center of mass for the entire system.[8] For a molecular system, E_ν is often represented by the overall kinetic energy due to the translational, vibrational, and rotational motions of individual molecules, the potential energy due to intra- and inter-molecular interactions, and the external energy arising from the molecular interactions with an external field (e.g., an electrical, gravitational, or magnetic potentials).

An ensemble average can also be defined in terms of microscopic quantities. Box 1.1 illustrates the ensemble average of an important microscopic quantity, the probability density of finding a particle at a given position. For a uniform system, this average is trivially known, $\rho = N/V$, where N is the number of particles and V is the system volume. When the system is inhomogeneous

---

[7] In statistics, $\langle M \rangle$ is called expectation, i.e., the expected value of M_ν.

[8] If the entire system is in motion relative to some frame of reference, the energy associated with that motion is not included in internal energy. Neither does the internal energy include the gravitational potential related to the elevation of the entire system relative to some frame of reference. While not considered here, it is possible to include macroscopic motion and gravity in statistical thermodynamics.

(e.g., particle distribution under the influence of an external field), the average probability density is known as the density function, which is a fundamental quantity in classical density functional theory (cDFT),[9] an advanced statistical-mechanical method to predict the equilibrium properties of thermodynamic systems.

Box 1.1 Density Function

Consider a thermodynamic system containing N identical particles. Assume these particles are classical such that the microstates of the system can be represented by the particle positions and momenta. At each microstate, the probability density of finding a particle is specified by the Dirac-delta function (see Figure 1.4 for a one-dimensional illustration). The instantaneous local density at position $\mathbf{r}$ is given by the summation of the probabilities of all particles

$$\hat{\rho}(\mathbf{r}) = \sum_{i=1}^{N} \delta(\mathbf{r} - \mathbf{r}_i).$$

The density function is defined as the ensemble average of the instantaneous local particle density

$$\rho(\mathbf{r}) = \langle \hat{\rho}(\mathbf{r}) \rangle.$$

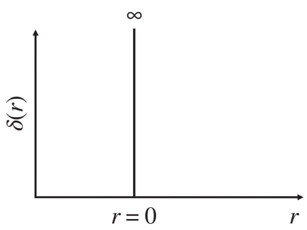
For a uniform system, the density function is invariant with position, i.e., the particle density is a constant, N/V.

1.3.4 Ergodicity

How do we know that the properties obtained from ensemble averages correspond to those measured in a particular thermodynamic system under study? What are the connections between different systems in the ensemble? Are the instantaneous microstates of the virtual copies of the thermodynamic system under study relevant to that of the real system? To answer these questions, we need the ergodic hypothesis, one of the most fundamental postulates of statistical mechanics.[10]

To understand ergodicity, imagine how the microscopic properties of a particular thermodynamic system vary with microstates. As shown schematically in Figure 1.5, the instantaneous value for any measurable thermodynamic quantity M changes with time because the microstate evolves with the motions of individual particles. If M is measured by some instrument, the experimental

Figure 1.4 The Dirac delta function $\delta(r)$ is everywhere zero except ∞ at $r = 0$. Because $\int \delta(r)dr = 1$, $\delta(r)$ may be understood as a special form of the probability density with the distribution of a continuous variable r localized to a single point.



---

[9] In mathematics, a functional maps a function (or functions) into a number, i.e., a functional is a function with function(s) as the input.

[10] There have been many attempts to prove the ergodic hypothesis from the dynamic perspective but a generic proof is yet to be established. For recent developments, see Moore C. C., *PNAS* 112 (7), 1907–1911 (2015).

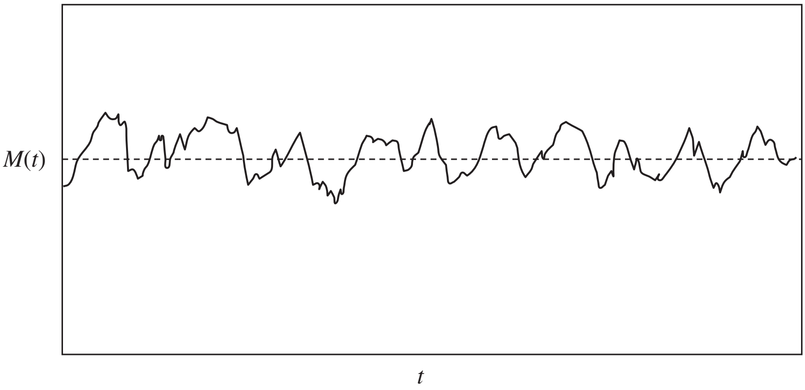


Figure 1.5 Schematic of time variance of dynamic property M of an equilibrium system (e.g., $M(t)$ could be the instantaneous total energy of a tank of argon gas at fixed N, V, and T). While M changes with time, at equilibrium, time average $\overline{M}$ (dashed line) is constant.

result corresponds to the time average of its instantaneous value over duration τ

$$\overline{M} = \frac{1}{\tau}\int_0^{\tau} M(t)\,dt. \tag{1.31}$$

Because thermodynamic quantities are time-invariant, the duration of measurement should be much larger than the time scale relevant to the dynamics of the individual particles. Besides, the experimental result should be independent of the time of measurement.

For a typical molecular system, the time scale relevant to the dynamics of individual atoms is exceedingly small, about 10^{-15} seconds or femtoseconds. Over the duration of a typical measurement for the equilibrium properties of a thermodynamic system (of the order of minutes or hours), the system is allowed to visit virtually all possible microstates. As a result, Eq. (1.31) can be equivalently written as

$$\overline{M} = \lim_{\tau\to\infty}\frac{1}{\tau}\int_0^{\tau} M(t)\,dt. \tag{1.32}$$

As the measurement samples all possible microstates, it is reasonable to conjecture that the microstates distribution is the same as that in the ensemble. In other words, the ensemble average and the time average would yield identical results

$$\langle M\rangle = \lim_{\tau\to\infty}\frac{1}{\tau}\int_0^{\tau} M(t)\,dt. \tag{1.33}$$

The equivalence of time average and ensemble average as given by Eq. (1.33) is called *ergodicity*.[11] Whereas Eq. (1.33) may not be valid for glassy systems that have multiple regions of microstates not mutually accessible to each other, we conjecture that the ergodic hypothesis holds true for all thermodynamic systems at equilibrium. Systems that do not satisfy ergodicity are referred to as non-ergodic.

We may illustrate ergodicity and the conceptual difference between time average and ensemble average by using an analogy. Suppose that we have two dice, each with six facets labeled with numbers 1–6. If we roll the two dice once, we want to know: What is the probability that the face values of the two dice give a seven? We can answer this question in two ways. In the first way, we have one

---

[11] The term “ergodic” was introduced originally by Ludwig Boltzmann based on the Greek words $\varepsilon'\rho\gamma o\nu$ (ergon: “work”) and $o\delta\acute{o}\varsigma$ (odos: “path” or “way”).

person roll the dice repeatedly, say n times. For each roll, we record the sum of the face values of the dice. For two different dice, there are 36 different outcomes (viz., "microstates"). We can calculate the desired probability by noting how often the two dice give seven and dividing that by n, the total number of rolls. This would give a time-average result because it comes from many rolls, each with the same dice at a different time. Alternatively, we can assemble many people and give each one two dice. The dice given to any one person are identical to those given to any other person. There are n persons in the assembly. At a fixed instant, everyone in the assembly rolls his or her dice. We can then calculate the desired probability by looking at all the rolls and counting the number of rolls that result in seven. We take that number and divide it by n, the number of persons who, at the same time, rolled his or her two dice. In this case, the calculated probability is not the average of many rolls taken by one person over an extended period. Instead, it is the average of many rolls, each taken by one of many persons (an ensemble) at a fixed time. We expect that the two methods will give the same result. Of the 36 possible outcomes, six outcomes give seven; they are: (1,6), (6,1), (2,5), (5,2), (3,4), (4,3). The desired probability is $6/36 = 1/6$.

1.3.5 Summary

In this section, we have discussed microscopic uncertainty as an intrinsic feature of all thermodynamic systems. Because of the microscopic uncertainty, we use statistical ensembles to define the probability distribution for microstates. A fundamental hypothesis in the ensemble approach is ergodicity, which asserts that properties evaluated from ensemble averages are equivalent to corresponding variables in time average as measured in a specific thermodynamic system.

1.4 Entropy and Information

As discussed in Section 1.1, thermodynamics is centered around two fundamental quantities, internal energy and entropy, from which other thermodynamic properties can be derived. While internal energy is linked with the kinetic and potential energies of individual particles, entropy provides a measure of the system *information*, i.e., diverse ways how individual particles may exist. In classical thermodynamics, entropy is often perceived as mysterious because, unlike internal energy, information does not have any obvious meaning without considering the uncertainties in the microscopic constituents of a thermodynamics system. In this section, we introduce two common definitions of entropy in statistical mechanics and explore different interpretations from microscopic perspectives.

1.4.1 The Boltzmann Entropy

In searching for a mechanical theory of heat, Ludwig Boltzmann concluded in 1872 that entropy is an intrinsic property of an equilibrium system reflecting the number of ways that individual particles may exist. According to Boltzmann, entropy is defined as

$$S = k_B \ln W \tag{1.34}$$

where $k_B = 1.381 \times 10^{-23}$ J/K is the Boltzmann constant,[12] and W stands for the number of *accessible* microstates, i.e., microstates that the system may take with nonzero probability. Boltzmann's definition of entropy was carved on his gravestone.

---

[12] Although Boltzmann first linked entropy and probability, the specific constant was introduced by Max Planck in the quantum theory of black-body radiation (see Section 4.3).

Eq. (1.34) provides a simple interpretation of entropy from a microscopic perspective: entropy is related to the number of choices available to the individual particles in a thermodynamic system. Because the particles may exist in different microstates, entropy provides a measure of the microscopic degrees of freedom.

We may illustrate the meaning of the total number of microstates by considering a simple example. Suppose that we have a system containing four *distinguishable* boxes, for example, by their fixed positions. We have four balls that can be randomly placed into these boxes, and we ask: In how many ways (microstates) can we place these four balls into the four boxes such that there is one ball in each box?

First, consider that all balls are different. When we place the first ball, we have four possibilities because all four boxes are initially empty. When we place the second ball, we have three possibilities because one box is already occupied. When we place the third ball, we have two possibilities and when we place the fourth ball, we have only one possibility. Therefore, the number of ways to place the balls is

$$W = 4 \cdot 3 \cdot 2 \cdot 1 = 24. \tag{1.35}$$

If the four balls are identical, i.e., indistinguishable, however, the final arrangements of balls in the four boxes are equivalent, independent of the sequence of filling. In that case, it makes no difference which ball we place first, second, third, or fourth. Therefore, we must divide the apparent W by 4!, giving $W = 1$. In other words, there is only one way to place four identical balls into four different boxes such that one ball is in each box. For identical balls, the number of microstates is unity. In these two scenarios, we say that the entropy for the case with different balls is higher than that with identical balls because the former provides more ways to arrange the balls (viz., more microstates).

But now suppose we have four balls where two are black and two are white; otherwise, the balls are identical. Again, we ask, in how many ways (microstates) can we place the four balls into the four boxes? For this case, we must divide the apparent W, as given by Eq. (1.35), by $2! \times 2!$, because two black balls and two white balls are indistinguishable. For this case, $W = 6$.

Figure 1.6 shows the six ways of filling four different boxes with two identical black balls and two identical white balls such that one ball is in each box. Because entropy is related to the number of microstates, the entropy of two identical black balls and two identical white balls in four different boxes is larger than that of four identical white balls (or four identical black balls) in four boxes.

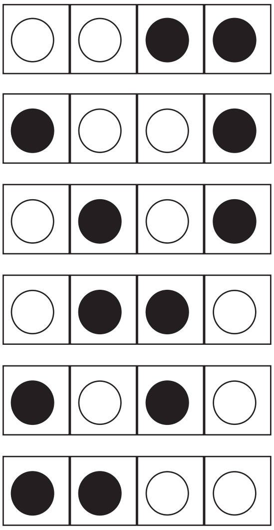


Figure 1.6 Six ways to place two identical white balls and two identical black balls into four different boxes such that one ball is in each box.

Why is it larger? When filling the four boxes with four identical balls, the information is complete; there is only one arrangement. But when filling four boxes with two white balls and two black balls, the information is incomplete. As shown in Figure 1.6, that case has six arrangements. We know that there are two black balls and two white balls in four boxes, but we do not know the detailed arrangement; information concerning the state of the system is incomplete.

The above example illustrates that, according to Boltzmann, entropy is a measure of the number of ways that a set of particles can be arranged. Here, different arrangements are affiliated with the particle identity. When the particles can be rearranged or shuffled among different configurations over time, the entropy of a system with the same balls is smaller than that of a system with different balls because identical balls reduce the number of ways that the balls can be arranged in the boxes. When otherwise identical balls are labeled, the entropy increases because the particles can be arranged in more different ways. In Box 1.2, we provide a more sophisticated example illustrating how entropy is intrinsically connected with such microscopic information.

Box 1.2 Maxwell's Demon

An imaginary being (a.k.a., a demon) was discussed in a letter written by James Clerk Maxwell to his life-long friend Peter Guthrie Tait in 1867, suggesting that the second law of thermodynamics would be violated. As illustrated in Figure 1.7, Maxwell conceived two gas chambers A and B with a shutter in between. An intelligent demon, who could keep track of the speeds of gas molecules, would be able to control the shutter such that high-speed molecules pass in one direction and low-speed molecules in the other. The biased passing of gas molecules through the shuttle would lead to temperature disparity between A and B. If the collisions of gas molecules with the shutter were elastic, and moving the shutter was frictionless, no work would be done by the demon. As a result, heat transfer takes place from elevated temperature to low temperature without any work.

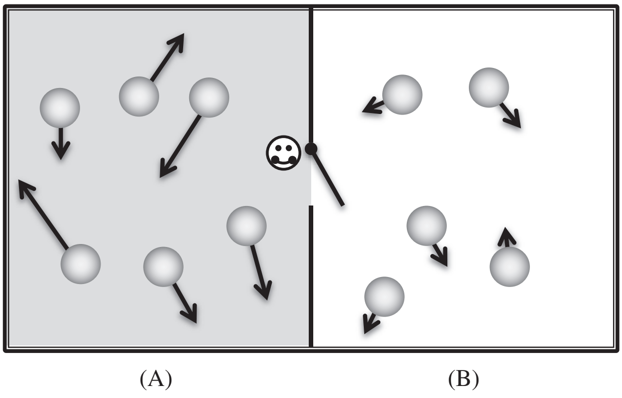


Figure 1.7 Maxwell's demon operated a trap door separating two volumes of gas A and B at initially equal temperature.

The demon's action reduces the total entropy of an isolated system (viz., gas molecules in A and B plus demon) thereby violating the second law of thermodynamics. The second law is saved, even if Maxwell's demon would be invented, only if one recognizes that the entropy decrease is achieved by the acquisition of microscopic information (here, the speed of individual molecules) by the intelligent demon. As illustrated in Box 1.3, Landauer's principle provides an explicit connection between information and thermodynamic quantities.

1.4.2 The Gibbs Entropy

The connection between entropy and microscopic information (more precisely, missing information) was first identified by J.W. Gibbs who provided an alternative definition of entropy[13]

$$S = -k_B \sum_{\nu} p_\nu \ln p_\nu. \tag{1.36}$$

As discussed in Section 1.3, p_ν stands for the probability of a thermodynamic system in microstate ν, and the summation extends to all accessible microstates.

In the context of ensemble average, Eq. (1.36) can be written as

$$S = -k_B \langle \ln p_\nu \rangle. \tag{1.37}$$

Eq. (1.37) suggests that, up to a universal constant k_B (which is needed for unit conversions between energy and temperature), the microscopic counterpart of entropy, $-\ln p_\nu$, can be understood as a measure of the degree of "uncertainty" or "incomplete information". When $p_\nu = 1$, the thermodynamic system has only one microstate ν. In this case, $-k_B \ln p_\nu = 0$ and $S = 0$ means no "uncertainty" at the microscopic level. For $p_\nu < 1$, the microstate of the system is not certain; the degree of "uncertainty" may be quantified in terms of $-\ln p_\nu$; the smaller p_ν is, the larger the system has "uncertainty".[14]

Box 1.3 Landauer's Principle

Rolf Landauer, an IBM scientist who made important contributions to information processing, analyzed the thermodynamics of data processing in 1961.[15] He discovered that information must be embodied in physical states and thus subject to thermodynamic regulations. For any computer operation that manipulates information, e.g., erasing a bit of memory, the entropy will increase due to loss of information about the physical state of individual logical units. The loss of information is manifested as the dissipation of heat accompanying the computer operation.

Landauer's principle states that for any irreversible single-bit operation on a physical memory element in contact with a heat bath at a given temperature, at least $k_B T \ln 2$ of heat must be released from the memory device into environment.

1.4.3 Connection Between Boltzmann's Entropy and Gibbs' Entropy

To establish a connection between the definition of entropy by Boltzmann and that by Gibbs, consider an equilibrium system that may adopt many microstates. Suppose that we can monitor the microstates of the system in $\mathbb{N}$ consecutive steps as time evolves, with $\mathbb{N}$ being a number that can be arbitrarily large. We designate the total number of microstates for the *entire* system as n. Within the duration of observation, the multiplicity of the system in different microstates is equal to the number of different outcomes to toss $\mathbb{N}$ times "a super die" with n facets

$$W_{\mathbb{N}} = \frac{\mathbb{N}!}{\mathbb{N}_1!\mathbb{N}_2!\ldots\mathbb{N}_n!} \tag{1.38}$$

---

[13] Gibbs' definition of entropy was later adopted by Claude Shannon in his theory of signal processing and data analysis, which constitutes an important cornerstone of information theory.

[14] Conceptually it is preferable to use a quantity that rises (rather than falls) as information increases. For that purpose, negentropy is defined as the reciprocal of entropy.

[15] Landauer R., "Irreversibility and heat generation in the computing process", *IBM J. Res. Develop.* 5 (3), 183–191 (1961).

where $\mathbb{N}_\nu$ is the number of times that the system is in microstate ν, satisfying the normalization condition

$$\sum_{\nu=1}^{n} \mathbb{N}_\nu = \mathbb{N}. \tag{1.39}$$

Because the multiplicity of the outcomes involves $\mathbb{N}$ steps, the average number of microstates at each step is

$$W = W_{\mathbb{N}}^{1/\mathbb{N}}. \tag{1.40}$$

In a special case where all microstates are equally accessible (e.g., a fair super coin), $W_{\mathbb{N}} = n^{\mathbb{N}}$, i.e., the number of accessible microstates at each step is n. In general, evaluation of $W_{\mathbb{N}}^{1/\mathbb{N}}$ requires some specific knowledge of the statistical distribution of the microstates.

According to Boltzmann, the entropy of the system is given by

$$S = k_B \ln W = \frac{k_B}{\mathbb{N}} \ln W_{\mathbb{N}}. \tag{1.41}$$

Using Eq. (1.38) for $W_{\mathbb{N}}$ and Stirling's approximation, $\ln \mathbb{N}! \approx \mathbb{N} \ln \mathbb{N} - \mathbb{N}$, we find that Boltzmann's definition of entropy is the same as that by Gibbs

$$\frac{S}{k_B} = \frac{\left(\ln \mathbb{N}! - \sum_{\nu=1}^{n} \ln \mathbb{N}_\nu!\right)}{\mathbb{N}} = -\sum_{\nu=1}^{n} (\mathbb{N}_\nu/\mathbb{N}) \ln(\mathbb{N}_\nu/\mathbb{N}) = -\sum_{\nu=1}^{n} p_\nu \ln p_\nu \tag{1.42}$$

where $p_\nu = \mathbb{N}_\nu/\mathbb{N}$ is the probability that the system is in microstate ν. Although in general the two definitions of entropy are not always identical (see Chapter 2), this simple example illustrates that both Gibbs' entropy and Boltzmann's entropy are related to the statistical distribution of microstates.

1.4.4 Entropy and Disorder

When thermodynamics is applied to diffusion or mixing processes, the entropy increase may be intuitively understood in terms of the rise of disorder or more randomness. However, a system in a state of higher entropy is not necessarily always equivalent to a state of greater disorder.[16] For certain systems, increasing entropy may serve as a driving force for microscopic ordering rather than disordering.

In the 1940s, Lars Onsager discovered that, at sufficiently high density, a system of rod-like particles exhibits a phase transition from an isotropic phase to a lamellar phase (much like a bunch of chopsticks, the rod-like particles are aligned in the axial direction).[17] The disorder-to-order transition can be explained by a simple thermodynamic model that accounts for the second-virial coefficient for interaction between rod-like particles. The phase transition occurs when the entropy of the ordered lamellar phase is higher than that of the disordered isotropic state. Entropy-driven ordering was also found by Alder and Wainwright in the 1950s using a computer model.[18] At sufficiently high density, spherical particles may crystallize without any attractive force (Figure 1.8). In other words, at sufficiently high density, the entropy of spherical particles in an ordered state may exceed that in a random state (see Section 7.9 for details).

---

[16] As commonly perceived, we define "order" in terms of spatial organization of individual elements. For example, a crystal is referred to as an ordered phase because of the organized arrangement of individual atoms.

[17] Onsager L., "The effects of shape on the interaction of colloidal particles", *Ann. NY Acad. Sci.* 51, 627–659 (1949).

[18] Alder B. and Wainwright T., "Phase transition for a hard sphere system", *J. Chem. Phys.* 5, 1208–1209 (1957).

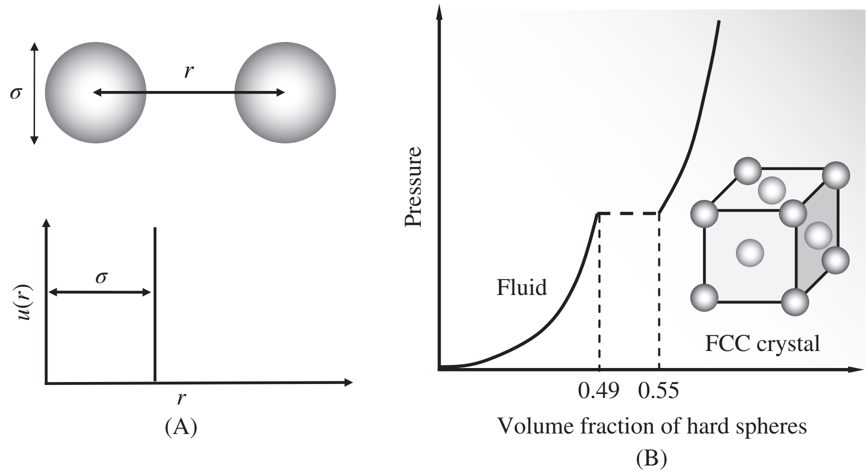


Figure 1.8 At sufficiently high volume fraction, a system of hard spheres maximizes its entropy by formating a face-centered cubic (FCC) structure. A. Schematic of hard-sphere potential between two identical particles; B. The pressure of a uniform hard-sphere system versus the volume fraction. The dashed lines mark the transition between disordered (random) and ordered (FCC) structures.

1.4.5 Entropy and Chaos

In classical thermodynamics, entropy is introduced through heat effects during reversible processes, i.e., idealized procedures for changing the thermodynamic conditions of a macroscopic system. Because the microscopic origin of heat is not intuitive, classical thermodynamics provides little insights on the physical significance of entropy. The lack of a physically comprehensible definition often results in misinterpretations. For example, many thermodynamics texts introduce entropy as a measure of disorder, randomness, or chaos; the higher the entropy, the greater the disorder. Such an interpretation may trace back to Rudolf Clausius who coined the term entropy from Greek, "en tropie," meaning transformation or change. Hermann von Helmholtz, another founding father of thermodynamics, also used German word "Unordnung" (disorder) to describe entropy. The interpretation is misleading because entropy may serve as a driving force of ordering. As discussed above and in more detail in Section 7.9, a thermodynamic state with a higher entropy can be more ordered than that with a lower entropy.

Rudolf Clausius proclaimed in 1865, "The entropy of the universe tends to a maximum." If entropy means disorder, the relentless increase of entropy would degrade our universe to a state of complete randomness or chaos. Historically, this bleak prediction of classical thermodynamics raised concerns among philosophers and laypersons alike. For example, Bertrand Russell wrote pessimistically that "all the labors of the ages, all the devotion, all the inspiration, all the noonday brightness of human genius, are destined to extinction."[19] Yet life emerges on earth, beginning from simple inorganic forms to biomacromolecules and their self-assembly to cells and eventual evolution to human beings with incremental ordered structures. Is entropy necessarily the same as disorder? Is "negative entropy" responsible for the organization in biological systems and the evolution of life?[20] Clarifying the concept of entropy with statistical mechanics will help answer these questions and avoid misunderstanding.

---

[19] Russell B., *The free man's worship*. Routledge, 1903.

[20] Erwin Schrödinger stated in his influential book, *What is Life*, that the emergence of life in the universe us due to the input of negative entropy.

1.4.6 Summary

According to statistical mechanics, entropy arises from uncertainties in microscopic details when a macroscopic system is defined by a few thermodynamic variables. As first recognized by Boltzmann, entropy provides a measure of the enormous number of microstates that a thermodynamic system may exist at any instant. Uncertainties in microscopic details are not necessarily equivalent to randomness or chaos – entropy can be a driving force of ordering in macroscopic systems. Mathematically, entropy is naturally affiliated with the information about individual particles in a thermodynamic system.

1.5 *Ab initio* Thermodynamics

When statistical mechanics was originally developed in the late nineteenth century, the subject was primarily an exercise of abstract mathematics. Many concepts were introduced without specifying the physical nature of the particles in thermodynamic systems or how they may interact with each other. Despite its importance, the precise meaning of particles or microstates was obscure. The dilemma was clearly stated in the classical text by J.W. Gibbs,[21] "*we avoid the gravest difficulties when, giving up the attempt to frame hypotheses concerning the constitution of material bodies, we pursue statistical inquiries as a branch of rational mechanics. In the present state of science, it seems hardly possible to frame a dynamic theory of molecular action which shall embrace the phenomena of thermodynamics, of radiation, and of the electrical manifestations which accompany the union of atoms. Yet any theory is obviously inadequate, which does not take account of all these phenomena. Certainly, one is building on an insecure foundation, who rests his work on hypotheses concerning the constitution of matter*."

Thanks to tremendous developments in theoretical physics, in particular, molecular sciences over the past century, Gibbs' view of statistical mechanics has been long outdated. In this and the next section, we will specify the microscopic details of thermodynamic systems from different perspectives: first principles, atomistic, coarse-grained, and lattice models. This section provides a very brief overview of the so-called first principles or *ab initio* thermodynamics, whereby microstates are defined in terms of the quantum states of electrons and atomic nuclei. All thermodynamic properties can be predicted from the statistical distribution of quantum states.

A detailed description of first-principles methods is beyond the scope of this book. The online Supplementary Material provides a brief introduction of the basic concepts that will be sufficient for understanding the materials discussed in this section. Because quantum states can be predicted without adjustable parameters, a combination of first-principles calculations with statistical mechanics empowers a quantitative description of both the micro- and macro-properties of molecular systems.

1.5.1 Quantum States

Quantum mechanics asserts that elementary particles of matter, such as electrons, protons, and photons, exit in enumerable microstates (viz., quantum states). Each quantum state can be represented by a wave function, and the properties of a many-body system are manifested as the expectation values of observables.

---

[21] Gibbs J. W., "*Elementary principles in statistical mechanics*", Dover Publications, reprint edition (2014).

While all microscopic properties can be described in terms of wave functions, exact quantum-mechanical methods are presently too complicated to be numerically soluble for thermodynamic systems of practical interest. Two approximations are commonly adopted in applications of quantum-mechanical methods to materials and chemical systems. First, the Born-Oppenheimer approximation assumes that the electronic degrees of freedom can be decoupled from those corresponding to atomic nuclei such that the atomic nuclei behave like classical particles with negligible size. Here, the degrees of freedom are referred to a complete set of variables to describe the dynamics of individual particles. By classical, we mean that quantum effects such as uncertainty and tunneling are irrelevant in describing particle motions and energy. The second assumption is that, in an atomic system, the electronic properties can be represented by those corresponding to electrons at 0 K in an external field due to interactions with the atomic nuclei. With these two assumptions, the microstates of any atomic system (viz., chemicals and materials) are defined by the quantum states of electrons while the nuclei are treated as classical particles with exact positions and momenta at any moment.

For most atomic systems of practical interest, the Born-Oppenheimer assumption is justified because of the small volume of atomic nuclei and the negligible electron rest mass m_e. On the one hand, the size of nuclear particles is tiny in comparison with the length scale characterizing electrons: the former is about few femtometers (10^{-15} m) but the latter on the order of Angstrom (10^{-10} m). As first discovered by Ernest Rutherford in 1911, "an atom is mostly empty space in which electrons obit". On the other hand, the rest mass of a proton, the smallest nuclear particle, is more than a thousand times the electron rest mass ($m_p/m_e \approx 1836$). The enormous difference in rest mass implies that the electron motion is much faster than that of atomic nuclei, typically by several orders of magnitude. As a result, electrons may be considered in the ground state at any atomic configuration defined by the positions of atomic nuclei.

1.5.2 Density Functional Theory (DFT)

In many practical applications, quantum-mechanical calculations are based on the electronic density functional theory (DFT), a mathematical procedure originally developed by Walter Kohn and others in the 1960s.[22] Rather than finding many-body wave functions directly, DFT seeks to solve quantum-mechanical problems by mapping many-body systems to noninteracting references. In comparison with alternative quantum-mechanical methods, DFT is drastically more efficient from a computational perspective yet applicable to a wide range of materials and chemical systems with reasonable accuracy.

To define the quantum states of electrons in an atomic system, DFT calculations start with the Kohn-Sham (KS) equation[23]

$$\left[-\frac{\hbar^2}{2m_e}\nabla^2 + v^{KS}(\mathbf{r})\right]\psi_\nu(\mathbf{r}) = \varepsilon_\nu \psi_\nu(\mathbf{r}) \tag{1.43}$$

where $\psi_\nu(\mathbf{r})$ stands for the wave function of noninteracting electrons, $\hbar = h/2\pi$, $h = 6.626 \times 10^{-34}$ J s is the Planck constant, $m_e = 9.109 \times 10^{-31}$ kg is the electron rest mass, $v^{KS}(\mathbf{r})$ represents the KS potential for an electron at position $\mathbf{r}$, and ε_ν stands for the electronic energy at a single-particle

---

[22] Hohenberg P. and Kohn W., "Inhomogeneous electron gas", *Phys. Rev.* 136 (3B), B864–B871 (1964); Kohn, W. and Sham, L. J., "Self-consistent equations including exchange and correlation effects", *Phys. Rev.* 140 (4A), A1133–A1138 (1965).

[23] An alternative to KS-DFT is provided by orbital-free DFT, which is numerically simpler but theoretically more challenging in formulation of an accurate functional.

quantum state ν. Mathematically, Eq. (1.43) resembles the one-particle Schrödinger equation, the simplest possible quantum-mechanical description of electronic systems.

In principle, DFT provides an exact mathematical framework to describe electronic properties at the ground state. Regrettably, the KS potential is only partially known. It includes contributions due to electron-nuclei interactions $v_N(\mathbf{r})$, the electrostatic repulsion among the electrons (viz., the Hartree energy) $v_H(\mathbf{r})$, and an exchange-correlation (xc) potential $v_{xc}(\mathbf{r})$

$$v^{KS}(\mathbf{r}) = v_N(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}). \tag{1.44}$$

The first two contributions on the right side of Eq. (1.44) are simply described by Coulomb's law, and the third term, $v_{xc}(\mathbf{r})$, is associated with the multi-body coupling of the Pauli exclusion principle (viz., exchange) and electrostatic interactions (viz., correlations). Because exact formulation of the exchange-correlation potential amounts to solving the many-body Schrödinger equation, approximations are inevitable in practical applications of KS-DFT. Over the years, numerical functionals have been developed to approximate $v_{xc}(\mathbf{r})$ and many software packages are available.[24] In general, the best choice of an xc-potential is dependent on the specific system and/or a given task, more like art than science.

1.5.3 Quantum Molecular Dynamics Simulation (QMD)

One major advantage of the first-principles approach is that it captures chemical reactions as well as intermolecular forces on an equal footing. More important, atomic motions, subsequently the evolution of microstates, can be naturally described by the fundamental laws of physics such that thermodynamic properties can be predicted by averaging over many microstates. The numerical procedure to integrate quantum and statistical mechanical calculations together is known as quantum molecular dynamics simulation (QMD) or *ab initio* molecular dynamics (AIMD). As the computer power ever increases, QMD/AIMD becomes increasingly more feasible and accurate in practical applications.

According to the Born-Oppenheimer assumption, each nuclear particle n experiences a Coulomb force due to interactions with other nucleic particles and with electrons in the surrounding

$$\mathbf{F}_n(\mathbf{r}) = \mathbf{F}_{nm}(\mathbf{r}) - \int d\mathbf{r}' \rho_e(\mathbf{r}') \frac{Z_n e^2(\boldsymbol{r} - \mathbf{r}')}{4\pi\varepsilon_0 |\mathbf{r} - \mathbf{r}'|^3} \tag{1.45}$$

where $\rho_e(\mathbf{r})$ represents the local electronic density, Z_n stands for the nuclear valence, $e = 1.602 \times 10^{-19}$ C is the elementary charge, and $\varepsilon_0 = 8.854 \times 10^{-12}$ J^{-1} C^2 m^{-1} is the dielectric permittivity in a vacuum. In Eq. (1.45), $\mathbf{F}_{nm}(\mathbf{r})$ represents the electrostatic force for the interaction of the nuclear particle with other nuclear charges

$$\mathbf{F}_{nm}(\mathbf{r}) = \sum_m \frac{Z_n Z_m e^2(\boldsymbol{r} - \boldsymbol{r}_m)}{4\pi\varepsilon_0 |\mathbf{r} - \mathbf{r}_{\mathbf{m}}|^3} \tag{1.46}$$

where the summation applies to all other nuclear particles in the system. The electronic density can be calculated from

$$\rho(\mathbf{r}) = \sum_{i=1}^{N_e} |\psi_i(\mathbf{r})|^2 \tag{1.47}$$

---

[24] Mardirossian N. and Head-Gordon M., "Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals", *Mol. Phys.* 115, (19), 2315–2372 (2017).

where N_e stands for the total number electrons in the system, and $\psi_i(\mathbf{r})$ is the single-electron wave function obtained from KS-DFT. With the force on each nuclear particle calculated from Eq. (1.45), the atomic motion can be described with the Newton equation

$$\mathbf{F}_n = m_n \mathbf{a}_n \tag{1.48}$$

where m_n stands for the nucleic mass, and $\mathbf{a}_n$ represents the particle acceleration. Once the atomic configuration is changed by numerically integrating the Newton equation, we update the electronic wave functions and, subsequently, the electronic density profile and the momenta of nuclear particles. The QMD/AIMD iteration leads to many microstates from which thermodynamic properties can be calculated from ensemble averages. In Chapter 2, we will discuss the mathematical details to integrate Newton's equation of motion for many particles.

1.5.4 Summary

In practical applications of statistical mechanics, we need an explicit definition of microstates. Otherwise, statistical mechanics would exist merely as an abstract formality, and the theoretical predictions could hardly be comparable with experimental measurements. In this section, we have discussed how "particles", i.e., the individual elements of thermodynamic systems, can be specified from first principles. Based on the Born-Oppenheimer assumption, we can define microstates in terms of the electronic quantum states. Accordingly, thermodynamic properties of an atomic system can be predicted from KS-DFT in combination with molecular dynamics simulation. The so-called *ab initio* thermodynamic procedure will be discussed in detail in Chapter 3 for ideal-gas systems.

1.6 Statistical-Thermodynamic Models

While tremendous progress has been made over the past few decades to predict the properties of materials and chemical systems from first principles, their applications to practical systems remain limited not only because of the computational burden but also because of insufficient numerical accuracy compromised by approximations introduced in quantum-mechanical calculations. As a result, semi-empirical models are commonly used in statistical thermodynamics to represent molecules or individual particles of a thermodynamic system from a classical perspective. In typical semi-empirical models, the microstates are often defined in terms of the positions and momenta of individual particles. The model parameters for describing particle-particle interactions can be calibrated with microscopic and/or macroscopic properties of thermodynamic systems. Based on limited experimental data and/or results from first-principles calculations, the semi-empirical models can predict thermodynamic properties of realistic chemical systems in good agreement with experimental measurements.

In this section, we discuss several types of statistical-thermodynamic models that are commonly employed for representing macroscopic systems, i.e., atomic, solvent-free, and coarse-grained models for materials and molecular systems, and lattice models that are often used to describe cooperative phenomena and universality of phase transition (Chapter 5).

1.6.1 Atomic Models

The materials world is made of a little over 100 types of atomic elements. However, natural phenomena are diverse and fascinating because atoms interact with each other through chemical bonds and longer-ranged forces. In statistical mechanics, atomic models intend to capture bond connectivity and intermolecular interactions, often by treating each atom as a classical particle. The microstates of an atomic system are thus represented by the positions and momenta of classical particles within the framework of Newtonian mechanics.

In a typical atomic model, the total energy of a thermodynamic system includes a kinetic energy (K) associated with the particle motions, and a potential energy (Φ) due to atomic interactions

$$E(\mathbf{r}^N, \mathbf{p}^N) = K(\mathbf{p}^N) + \Phi(\mathbf{r}^N) \tag{1.49}$$

where N represents the total number of atoms, $\mathbf{r}^N \equiv (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ is the atomic configuration as defined by the positions of all atoms, and $\mathbf{p}^N \equiv (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N)$ represents the momenta of all the atoms. The multidimensional space spanned by the composite vector, $\mathbf{x} \equiv (\mathbf{r}^N, \mathbf{p}^N)$, is called the *phase space*, which defines the degrees of freedom of all particles in the system.

When the atoms are depicted by a classical model, the total kinetic energy is simply determined by the atomic mass (m_i) and momenta ($\mathbf{p}_i$) for each particle

$$K(\mathbf{p}^N) = \sum_{i=1}^{N} |\mathbf{p}_i|^2/2m_i. \tag{1.50}$$

Meanwhile, the potential energy consists of contributions resulting from chemical bonding (Φ^B) and nonbonded (Φ^{NB}) interactions[25]

$$\Phi(\mathbf{r}^N) = \Phi^B(\mathbf{r}^N) + \Phi^{NB}(\mathbf{r}^N). \tag{1.51}$$

Over the years, many semi-empirical functions have been developed to describe the potential energy of atoms in materials and molecular systems. These functions, along with parameters designed to reproduce molecular structure and selected properties of thermodynamic systems, are commonly referred to as *molecular force fields*.

With an analytical expression for the potential energy, we can evaluate the net force on each atom based on its relative position with all other atoms in the system

$$\mathbf{F}_i(\mathbf{r}^N) = -\nabla_{\mathbf{r}_i}\Phi(\mathbf{r}^N). \tag{1.52}$$

Based on an initial condition specifying the positions and momenta of all particles, Eq. (1.52) can be used to predict the evolution of atomic positions (viz., a *trajectory of atomic motions*) by integrating Newton's equation. The numerical procedure is known as classical molecular dynamics simulation, or MD, which will be discussed in Chapter 2.

In comparison with QMD or AIMD, classical MD simulation is computationally less demanding due to the absence of quantum-mechanical calculations. Nevertheless, a supercomputer is often needed for MD simulation of realistic systems because of the system size and time scales. The time scale pertinent to atomic motions is in the order femtoseconds while the diameter of an atom is only a few Ångstrom ($1\ \text{Å} = 10^{-10}$ m). The computational cost can be drastically reduced by using coarse-grained models where each particle is ascribed to a group of atoms (e.g., functional groups or monomeric segments of a polymer) or molecules.

1.6.2 The Lennard-Jones Model

To gain an understanding of how the microscopic details of individual particles determine the macroscopic properties of a thermodynamic system, it is useful to consider simple models. One

---

[25] Supplementary Materials IV.

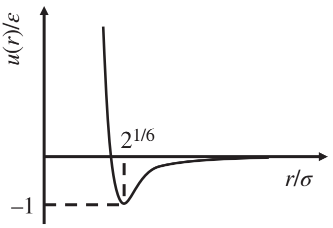


Figure 1.9 In dimensionless units defined by the energy and size parameters of the Lennard-Jones model, the intermolecular potential is a universal function of the distance.

of the simplest atomic models for molecular systems assumes that each molecule is a spherical particle, much like a billiard ball with attraction. The model provides a good description of the thermodynamic properties and phase behavior of simple fluids such as noble gases and systems consisting of small nonpolar molecules (e.g., methane and nitrogen).

In a simple fluid, the intermolecular interaction is often represented by the Lennard-Jones (LJ) potential

$$u(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{1.53}$$

where r is the center-to-center distance, parameters σ and ε are related to the particle size and the strength of maximum attraction. The second term on the right side of Eq. (1.53) represents the van der Waals attraction, and the first term approximates the short-range repulsion. A similar equation can be written for the pair potential between different types of small nonpolar molecules.

As shown in Figure 1.9, the LJ potential suggests a universal form of intermolecular interaction when the distance and potential energy are normalized by parameters σ and ε, respectively. The universal form offers a statistical-mechanical basis for understanding the theorem of corresponding states first proposed by van der Waals, i.e., in dimensionless variables of k_BT/ε and $\rho\sigma^3$, the thermodynamic properties of a simple fluid can be represented by universal functions of two independent variables as predicted by the Gibbs phase rule.[26]

Tremendous research effort has been devoted to LJ systems because once such systems are understood, it is then possible to apply corrections (viz., perturbations) toward representing chemical systems with more complicated intra- and inter-molecular interactions. This procedure is often used in classical thermodynamics. For example, we use the well-understood properties of an ideal gas as a basis for understanding real gases; we apply corrections through a compressibility factor that can be obtained from a corresponding state correlation or from a realistic equation of state. Similarly, for vapor–liquid equilibria of nonelectrolyte liquid mixtures, we start with an ideal mixture (Raoult's law) and then apply corrections through excess Gibbs energy models for activity coefficients. We will discuss in Chapters 7–9 that simple fluids play a pivotal role in the development of liquid-state theories and equations of state for bulk chemical systems based on statistical mechanics.

1.6.3 Implicit-Solvent Models

In solution thermodynamics, the solute behavior is often of primary interest for both practical and fundamental reasons. For convenience and physical clarity, it is useful to formulate microstates in terms of solute molecules without an explicit consideration of the microscopic details of the solvent. The basic idea is akin to neglecting the electronic degrees of freedom in modeling thermodynamic systems with many atoms. As the electronic effects can be described in terms of molecular structure

---

[26] See, e.g., Tester, J. W. and Modell, M., *Thermodynamics and its applications* Prentice Hall, 1997, p. 244.

and semi-empirical force fields, the solvent effects can be incorporated into particle solvation and the solvent-mediated interactions. Because of their simplicity, the implicit solvent models have been widely adopted in theoretical descriptions of electrochemical systems, aqueous solutions, and colloidal dispersions.

By treating the solvent as a continuous medium, we can define the microstates of a solution or colloidal dispersion in terms of the positions and momenta of individual solute particles, similar to those in a solvent-free system. The particle motions in a solvent are subject to the friction generated by the solvent molecules as well as random forces due to particle collision with the solvent molecules. For historical reasons, the dynamics of particle motion in a solvent is called Brownian motion.[27] Because of the particle-solvent interactions, the total energy of the particles in Brownian motion is not conserved. Subsequently, the evolution of microstates is different from that for solvent-free systems. We will discuss Brownian motion in Section 2.5 and the applications of the solvent-explicit models to solution thermodynamics in Chapter 9.

1.6.4 Lattice Models

To capture the *universal* behavior of natural phenomena emerging from interactions among many particles, we desire an idealized model that retains the essential physics without attention to chemical details. Toward that end, lattice models are often adopted in statistical mechanics. In addition to their broad relevance to practical applications as discussed in Chapter 5, many fundamental concepts of statistical mechanics, such as mean-field approximations, correlations, and universality, are best understood within the framework lattice models.

Unlike electrons and classical particles discussed above, “particles” in a lattice model may not have a precise physical significance. In fact, here particles are simply represented by fundamental constituents of a macroscopic system, which can be constructed in different contexts. For example, in the lattice model, the lattice sites may be affiliated with gas molecules or polymer segments, “spin” orientations, particles with different identities, charge or energy states, etc. If each lattice site has only two states, the lattice model is commonly known as the Ising model, proposed originally by Ernst Ising to describe the magnetization of ferromagnetic materials but applicable to describe many cooperative phenomena in nature. In Chapter 5, we will discuss several applications of the Ising model by mapping the energy states of each lattice site with different physical analogies.

To elucidate the basic ideas, Figure 1.10 shows, for example, a thermodynamic system represented by a square lattice. Here, each site may be affiliated with one of two-energy states. A microstate of the system is specified by the energy state of each lattice site. Because the physical nature of the lattice sites is arbitrary, the model is useful to represent any two-dimensional system consisting of individual elements with two-energy states. For example, these lattice sites may be considered as the discretized positions or orientations of atoms, molecules, and the monomeric segments of polymer chains; they may also be assigned to the conformations of the building blocks of biomacromolecules or to the orientations of spins carried by elementary particles.

Intuitively, the two-dimensional lattice model may also be used to describe the monolayer absorption of gas molecules at a flat surface: the lattice may be identified as a discretized representation of the positions of the adsorbed molecules, and the states of each lattice site correspond to the position being occupied by a single gas molecule or empty. To define the microstates, we may assign

---

[27] Named after Robert Brown, an English Botanist who in 1827 first reported the chaotic motion of pollen grains suspended in liquid water. The observation had a pivotal role in the early development of modern atomic theory.

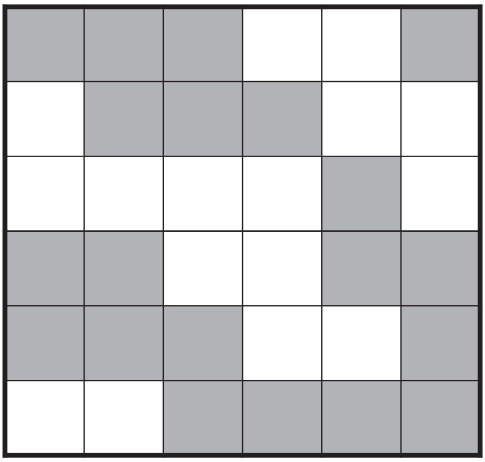


Figure 1.10 In a two-dimensional lattice model, each lattice site takes one of two-energy states as shown in different colors. The dichotomous states may represent, for example, the occupation status of a discretized surface by gas molecules or spin orientations.

any lattice site i with two integers, e.g.,

$$n_i = \begin{cases} 0 & \text{empty} \\ 1 & \text{occupied} \end{cases}. \tag{1.54}$$

Here, we use numbers (0,1) for simplicity; other choices, e.g., (−1,1), are also possible. For each microstate, we have a set of the numbers for $\{n_i\}$ to define the status of the lattice occupation. The system energy can be quantified by assigning each occupied lattice site with energy h, which accounts for the interaction between the surface and a gas molecule, and each unoccupied site with zero energy. An additional parameter, ϵ, may be used to describe nearest-neighbor interactions between gas molecules on the surface, i.e., an energy of ϵ is assigned for each pair of occupied sites directly in contact with each other.

At a given microstate, we have a set of numbers to describe the occupation status of each lattice site, i.e., $n_i = 0$ or 1, for $i = 1, 2, \ldots, N_s$, and N_s is the total number of lattice sites. The total energy at microstate ν can then be written as

$$E_\nu = -\sum_{i=1}^{N_s} n_i h - \frac{1}{2}\sum_{i=1}^{N_s}\sum_{j=1}^{N_n} n_i n_j \varepsilon \tag{1.55}$$

where h stands for a one-body energy, ε denotes a pair interaction energy, N_n represents the number of nearest neighbors for each site, and ½ accounts for the double counting of the nearest-neighbor pairs. In certain cases, we can find the probability distribution of the system in all microstates, p_ν. With this probability, we achieve a *statistical* description of the thermodynamic properties of the system.

We may evaluate ensemble averages using the two-dimensional lattice model for gas adsorption on a planar surface. At microstate ν, the occupation number for each lattice site takes the value of $n_{i,\nu} = 0$ or 1, depending on whether the site is covered by a gas molecule. The ensemble average of the occupation number represents the probability of this lattice site covered by a gas molecule

$$\langle n_i \rangle = \sum_\nu p_\nu n_{i,\nu}. \tag{1.56}$$

If all lattice sites are equivalent, at any moment, the total number of molecules on the surface is

$$N_\nu = \sum_{i=1}^{N_s} n_{i,\nu} \tag{1.57}$$

where N_s is the number of lattice sites. The surface coverage, which can be expressed as an ensemble average of N_ν

$$\theta = \frac{\langle N_\nu \rangle}{N_s} \tag{1.58}$$

is a thermodynamic quantity that can be measured by experiment.

1.6.5 Summary

Both physical and lattice models are commonly used in statistical mechanics. Physical models allow us to evaluate thermodynamic quantities associated with the various forms of kinetic and potential energies. If we are not concerned with dynamic or time-dependent behavior, physical models can be further simplified by ignoring particle motions while capturing only particle-particle interactions as represented by lattice models. The latter are useful to understand cooperative phenomena emerging from many-body interactions in thermodynamic systems. Lattice models are often used to describe the universal behavior of macroscopic systems such as those underlying phase transitions.

1.7 Additivity and Relativity in Thermodynamics

In this section, we discuss two important characteristics of internal energy and entropy, following their definitions in statistical mechanics. These characteristics help us better understand under what conditions extensive thermodynamic quantities are additive, why we can describe the properties of the same thermodynamic system using different statistical-mechanical models, and how we may compare theoretical predictions with experimental data.

1.7.1 Additivity of Extensive Variables

Consider a thermodynamic system consisting of two *independent* subsystems A and B that are macroscopically identical, i.e., the subsystems have the same composition and thermodynamic properties and are not interacting with each other.

As the subsystems are independent, the microstate of the composite system of A and B is defined by the microstates of both subsystems, i.e.,

$$\nu = (\nu_A, \ \nu_B). \tag{1.59}$$

At any instant, the microstate of subsystem A may not be the same as that of subsystem B. Because the composite system combines the degrees of freedom for both systems, the probability of microstate ν is

$$p_\nu = p_{\nu_A} \cdot p_{\nu_B} \tag{1.60}$$

where p_{ν_A} and p_{ν_B} are the microstate probabilities in subsystems A and B, respectively.

With the normalization conditions of microstate probabilities for both sub-systems

$$\sum_{\nu_A} p_{\nu_A} = 1, \tag{1.61}$$

$$\sum_{\nu_B} p_{\nu_B} = 1, \tag{1.62}$$

we can express the total entropy of the composite system as

$$\begin{aligned}\frac{S}{k_B} &= -\sum_{\nu} p_{\nu} \ln p_{\nu} = -\sum_{\nu_A}\sum_{\nu_B}(p_{\nu_A}\cdot p_{\nu_B})\ln(p_{\nu_A}\cdot p_{\nu_B})\\ &= -\sum_{\nu_B} p_{\nu_B}\sum_{\nu_A} p_{\nu_A}\ln p_{\nu_A} - \sum_{\nu_A} p_{\nu_A}\sum_{\nu_B} p_{\nu_B}\ln p_{\nu_B} = \frac{S_A}{k_B}+\frac{S_B}{k_B}\end{aligned} \tag{1.63}$$

Eq. (1.63) indicates that, as expected, entropy is additive. Similarly, the total internal energy is given by

$$\begin{aligned}U &= \sum_{\nu} p_{\nu}E_{\nu} = \sum_{\nu_A}\sum_{\nu_B}(p_{\nu_A}\cdot p_{\nu_B})[E_{\nu_A}+E_{\nu_B}]\\ &= \sum_{\nu_B} p_{\nu_B}\sum_{\nu_A} p_{\nu_A}E_{\nu_A} + \sum_{\nu_A} p_{\nu_A}\sum_{\nu_B} p_{\nu_B}E_{\nu_B} = U_A + U_B\end{aligned} \tag{1.64}$$

which is also additive.

Additivity holds true for all extensive thermodynamic variables. While for simplicity Eqs. (1.63) and (1.64) are obtained for identical subsystems, they are equally valid when subsystems A and B are not thermodynamically identical (e.g., A and B have different sizes or even different species). Additivity breaks down if the microstates of subsystems A and B are not independent. This would be the case, for example, if particles from subsystem A are interacting with those in subsystem B.

1.7.2 Relativity of Energy and Entropy

Physical models are indispensable in the statistical-mechanical description of thermodynamic properties. However, the selection of such models in practice appears subjective. We may take *ab initio* thermodynamics, atomic models, coarse-grained models, or lattice models for the same thermodynamic system. Are the thermodynamic quantities of a specific system not dependent on the definition of microstates? How do we compare results from different statistical-mechanical models (e.g., atomic against lattice models) for the same thermodynamic system? If the models are subjective, how are the theoretical results compared with experimental data? These questions can be answered by recognizing the relative nature of entropy and internal energy.

With different definitions of microstates, we have different numerical results for both the total number of microstates and the probability distributions. To elucidate their connections, suppose we first define the microstates of a thermodynamic system in terms of quantum states following the most fundamental principles of physics. In such an extreme approach, a microstate would contain information not directly relevant to thermodynamic properties of practical interest. For example, for a typical chemical system, the quantum states of sub-atomic particles within each atom are independent of those from other atoms. Therefore, for each atomic state as defined by, for example, the atomic configuration and kinetic energies, the sub-atomic quantum states may be considered as a subset of microstates. The summation over the microstates of both atomic and sub-atomic degrees of freedom is equivalent to the summation over the atomic states and that over the sub-atomic quantum states, i.e.,

$$\sum_{\nu} = \sum_{a}\sum_{n\in a} \tag{1.65}$$

where subscripts a and n denote atomic and sub-atomic states, respectively. The probability of the system in a particular atomic state is given by the summation over the probabilities of microstates with the sub-atomic states affiliated with the corresponding atomic state

$$p_a = \sum_{n \in a} p_{\{a,n\}} \tag{1.66}$$

where $p_{\{a,\, n\}} = p_\nu$ stands for the probability of the system in a particular microstate with sub-atomic details.

From the probabilities of microstates, we can write the entropy as

$$S = -k_B \sum_\nu p_\nu \ln p_\nu = -k_B \sum_a \sum_{n \in a} p_{\{a,n\}} \ln p_{\{a,n\}}, \tag{1.67}$$

and the internal energy is

$$U = \sum_a \sum_{n \in a} p_{\{a,n\}} E_{\{a,n\}}. \tag{1.68}$$

Using Eq. (1.66) and conditional probability

$$p_{n/a} = p_{\{a,n\}} / p_a, \tag{1.69}$$

we can rewrite the second summation on the right side of Eq. (1.67) as

$$\begin{aligned} \sum_{n \in a} p_{\{a,n\}} \ln p_{\{a,n\}} &= \sum_{n \in a} p_{\{a,n\}} \ln[p_a (p_{\{a,n\}} / p_a)] \\ &= p_a \ln p_a + p_a \sum_{n \in a} p_{n/a} \ln p_{n/a} \end{aligned} \tag{1.70}$$

Thus, the entropy can be decomposed into two terms

$$S = S_g + \sum_a p_a S_a \tag{1.71}$$

where subscript "g" and "a" stand for the grouped entropy and atomic entropy with

$$S_g = -k_B \sum_a p_a \ln p_a, \tag{1.72}$$

$$S_a = -k_B \sum_{n \in a} p_{n/a} \ln p_{n/a}. \tag{1.73}$$

Eq. (1.71) indicates that the entropy defined according to the microstates of sub-atomic particles is equal to that defined by the atomic states (viz., grouped microstates) plus an additional contribution associated with the sub-atomic degrees of freedom. The latter defines a subsystem entropy for each atom. If the subsystem entropy is a constant independent of the atomic state, the overall entropy becomes

$$S = S_g + S_a. \tag{1.74}$$

Eq. (1.74) suggests that entropy is relative, i.e., its absolute value depends on the relative degrees of freedom accounted for by the microstates.

Similarly, we can show that the total internal energy also satisfies the grouping property

$$U = U_g + U_a \tag{1.75}$$

where the two terms on the right represent the average atomic energy and the average energy of sub-atomic particles at each atomic configuration

$$U_g = \sum_a p_a E_a, \tag{1.76}$$

$$U_a = \sum_{n \in a} p_{n/a} E_{n/a}. \tag{1.77}$$

While for illustration the above discussion is concerned with the atomic and sub-atomic degrees of freedom, a similar procedure is applicable to other ways of grouping the microstates (e.g., in terms of clusters of atoms or lattice sites). Different coarse-grained models correspond to separate ways of grouping the microstates into coarse-grained particles. Except a model-dependent constant that cancels out in evaluating the changes of thermodynamic quantities, different statistical-mechanical models are equivalent because they should yield the same relative entropy, relative internal energy, and other relative thermodynamic quantities. As practical applications are concerned only with the changes of thermodynamic properties, different statistical-mechanical models are equivalent in principle; they differ only in grouping the microstates.

1.7.3 Summary

The internal energy and entropy of a thermodynamic system can be described as the sum of those corresponding to sub-systems. This additivity principle, which is also known as the extensiveness of thermodynamic quantities, is valid only when the sub-systems are independent. The relativity of thermodynamic quantities indicates that the variations of internal energy and entropy are model-independent. It should be noted that, in physics literature, the relativity of thermodynamic quantities may also refer to the fact that thermodynamic properties can be described in a frame-independent manner, regardless of the relative motion of the observer. The principle of relativity ensures that the thermodynamic laws are universally applicable, regardless of the reference frame used to describe the system.

1.8 Chapter Summary

This introductory chapter discusses some of the most basic concepts in describing a thermodynamic system from a microscopic perspective such as individual particles, microstates, ensemble, and ergodicity. These concepts allow us to better understand the physical meanings of internal energy and entropy – two most fundamental thermodynamic variables. From internal energy and entropy, we can derive auxiliary thermodynamic functions using the fundamental equations and Legendre transformations. Because both internal energy and entropy satisfy additivity and relativity principles, we may use complementary statistical-mechanical models to define the microstates of a thermodynamic system.

Statistical mechanics is applicable to materials and chemical systems where the dynamics of individual particles are described by fundamental laws of physics, as well as to, in general, many-body systems where particles may not have precise physical significance. In the former case, the microstates are often defined in terms of the degrees of freedom for realistic particles such as atoms or various “coarse-grained” representations of molecules or other basic constituents of a macroscopic system. While information concerning physical details is paramount for understanding the microscopic structure and thermodynamic properties of materials and chemical systems, abstract models are also important in statistical mechanics for their broad applicability to a wide variety of physical phenomena, ranging from cooperativity in many-body systems to critical behavior and the universality of phase transitions. Simple models such as a lattice representation of particle positions are important for pedagogical purposes, for example, to elucidate the basic concepts in statistical mechanics such as cooperativity, fluctuation phenomena, order parameters, phase transition, and broken symmetry. These fundamental concepts are universally applicable to systems containing many particles. The flexibility in choosing the microscopic ingredients of thermodynamic systems, either physical or abstract, makes statistical mechanics a versatile tool for understanding diverse phenomena in nature as well as chemical and material systems of practical interest.

Further Readings

Chandler D., *Introduction to modern statistical mechanics*. Oxford University Press Chapters 1–2, 1987.

Frenkel D., "Order through entropy", *Nat. Mater.* 14, 9–12 (2015).

Gubbins K. E. and Moore J. D., "Molecular modeling of matter: impact and prospects in engineering", *Ind. Eng. Chem. Res.* 49, 3026–3046 (2010).

Jaynes E. T., "Gibbs vs Boltzmann entropies", *Am. J. Phys.* 33, 391–398 (1965).

Moore G. C., "Ergodic theorem, ergodic theory, and statistical mechanics", *PNAS* 112 (7) 1907–1911 (2015).

Tsallis C., "Entropy", *Encyclopedia* 2 (1), 264–300 (2022).

Problems

1.1 Using the phase rule, justify whether the following thermodynamic systems are adequately defined:
- (i) a pure liquid of known mass at given temperature and density;
- (ii) a pure liquid of known mass at given density and pressure;
- (iii) a two-phase system at given temperature, pressure, and the mole fractions of all chemical species in one phase;
- (iv) an ionic liquid containing two types of cations and one type of anions at given temperature, pressure, and the ionic composition.

1.2 Based on the fundamental thermodynamic relations, show

$$\left(\frac{\partial C_P}{\partial P}\right)_{T,N} = -T\left(\frac{\partial^2 V}{\partial T^2}\right)_{P,N}.$$

1.3 Acoustic velocity, or the speed of sound, in a fluid can be derived from the first law of thermodynamics for steady-state flow processes

$$d(\hat{H} + v_s^2/2) = 0$$

in combination with the mass-balance equation

$$d(\hat{\rho} v_s) = 0.$$

Here, $\hat{H}$ is the specific enthalpy of the fluid, $\hat{\rho}$ is the mass density, and v_s is the velocity of the elastic wave, i.e., the sound speed.

(i) Derive the following expression from the 1st law and mass balance

$$v_s = \sqrt{\left(\frac{\partial P}{\partial \hat{\rho}}\right)_S}.$$

(ii) Show that the partial derivative at constant entropy S can be expressed in terms of measurable quantities

$$v_s = \sqrt{\frac{\gamma}{\hat{\rho}\kappa_T}},$$

where $\gamma \equiv C_P/C_V$ is the ratio of the constant-pressure and constant-volume heat capacities of the fluid, κ_T is the isothermal compressibility

$$\kappa_T \equiv -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T.$$

(iii) Predict the speed of sound in the atmosphere at 25 °C. Assuming atmosphere is an ideal gas with molar heat capacity $C_P = 29.3$ J/(mol K) and molecular weight $M = 29$ g/mol.

1.4 Consider a one-dimensional model for a thermodynamic system containing N identical particles. Assume that the internal energy is completely determined by temperature T, system total length L, and particle number N. Show that the fundamental equation of thermodynamics can be written as

$$dU = TdS + fdL + \mu dN,$$

where temperature, line tension (i.e., negative of the one-dimensional pressure), and chemical pressure are defined as, respectively,

$$T = \left(\frac{\partial U}{\partial S}\right)_{L,N},$$

$$f = \left(\frac{\partial U}{\partial L}\right)_{S,N},$$

$$\mu = \left(\frac{\partial U}{\partial N}\right)_{L,S}.$$

What are the analogs of the Maxwell relations for this one-dimensional system?

1.5 Two-dimensional models are often used to describe the thermodynamics of molecular adsorption from a gas or a liquid solution on planar surfaces. For a one-component two-dimensional system, the fundamental equation is given by

$$dU = TdS - \varsigma dA + \mu dN,$$

where ς stands for surface pressure. Derive the Gibbs adsorption isotherm:

$$\left(\frac{\partial \varsigma}{\partial \mu}\right)_T = \frac{N}{A}.$$

Can you extend the Gibbs adsorption isotherm to multicomponent systems? What is the surface pressure if there is no adsorption?

1.6 Consider gas adsorption on a planar surface at temperature T. Assume that the gas phase is ideal and that the adsorption can be described by Henry's law

$$N/A = k_H P,$$

where N is the number of gas molecules at the surface, A is the surface area, and k_H is Henry's constant. Show that the surface pressure is given by

$$\varsigma A = N k_B T,$$

where k_B is Boltzmann's constant.

1.7 Consider molecular adsorption from a liquid solution on a planar surface of area A and temperature T. Assume that the liquid phase is an ideal solution and that the adsorption of the solute molecules can be described by Henry's law

$$N/A = k_H C,$$

where N is the *moles* of solute molecules at the surface, k_H is Henry's constant, and C stands for the *molar* concentration of the solute molecules in the bulk solution. Show that the surface pressure is given by

$$\varsigma A = NRT,$$

where R is the gas constant.

1.8 Consider ideal-gas adsorption on a planar surface such that gas molecules at the surface follows a two-dimensional equation of state

$$\varsigma(a - a_0) = k_B T,$$

where a is the surface area per molecule, and a_0 is the surface area occupied by each molecule. Show that the adsorption isotherm follows the Volmer isotherm

$$Pk_H = \frac{\theta}{1-\theta} \exp\left(\frac{\theta}{1-\theta}\right),$$

where $\theta = a_0/a$ stands for the surface coverage, and k_H is Henry's constant defined by $k_H = \theta/P$ as $P \to 0$. How would you extend the Volmer equation to adsorption on a planar surface from an ideal solution?

1.9 Consider ideal-gas adsorption on a planar surface such that gas molecules at the surface follows a two-dimensional analog of the van der Waals equation of state

$$\varsigma = k_B T/(a - a_0) - c/a^2,$$

where a is the surface area per molecule, a_0 is the surface area occupied by each molecule, and c is an energy parameter that accounts for the intermolecular attraction at the surface. Show that the adsorption isotherm follows the Hill–Deboer equation

$$\ln\left[\frac{P(1-\theta)}{\theta}\right] - \frac{\theta}{1-\theta} = -\ln K_1 - \frac{K_2\theta}{k_B T},$$

where $\theta = a_0/a$ stands for the surface coverage, K_1 is defined by $K_1 = \theta/P$ as $P \to 0$ (viz., Henry's constant), and $K_2 = 2c/a_0$.

1.10 A container includes a large number of gas molecules that do not interact with each other. Assume that the container wall is elastic and that the gas molecules can be represented by non-interacting classical particles. Does the system satisfy the ergodic hypothesis? Why?

1.11 The Gibbs equation for entropy can be derived from the hypothesis that entropy is an ensemble average of some microscopic property depending only on the microstate probability, i.e.,

$$S = \sum_\nu p_\nu f(p_\nu),$$

where $f(p_\nu)$ may be understood as the microscopic counterpart of entropy at microstate ν. Show that this hypothesis leads to $f(p_\nu) = -k_B \ln p_\nu$. [Hint: Consider entropy additivity for independent thermodynamic systems.]

1.12 Suppose that a thermodynamic system has an initial distribution of microstates specified by $\{p_\nu^0\}$, where ν denotes microstates, and reaches a new microstate distribution $\{p_\nu\}$ at equilibrium. Show that

$$\sum_\nu p_\nu \ln p_\nu \geq \sum_\nu p_\nu \ln p_\nu^0,$$

where the equality holds if and only if $p_\nu = p_\nu^0$ for all ν. [Hint: $\ln x \leq x - 1$ for all $x > 0$ with equality if and only if $x = 1$.]

1.13 Consider a lattice model for a one-component gas such that each site can accommodate no more than one gas molecule. Assume that each microstate of the system is defined by a particular occupation of the lattice sites by the gas molecules.

(i) What is the number of microstates for the lattice system with n-sites containing N gas molecules?

(ii) Show that the lattice model predicts an entropy

$$S = -nk_B\left[x \ln x + (1-x)\ln(1-x)\right],$$

where $x = N/n$. [Hint: $\ln n! \approx n \ln n - n$.]

(iii) Assume that the system volume is proportional to the number of lattice sites, i.e., $V = nv_0$, where v_0 is the volume for each site. Show that, when $x \to 0$ (i.e., low gas density), the entropy change in response to the volume change from V_1 to V_2 is

$$\Delta S = Nk_B \ln(V_2/V_1).$$

(iv) Show that, when $x \to 0$, the lattice model satisfies the ideal-gas law $PV = Nk_BT$.

(v) How does the entropy change with temperature according to this model?

1.14 Consider a lattice model for a one-component gas such that each lattice site can accommodate no more than one gas molecule. Each gas molecule on the lattice is able to take m orientations with equal energy. Assume that the system volume is proportional to the number of lattice sites, $V = nv_0$, where n is the number of lattice sites, and v_0 is the volume per lattice site.

(i) What is the entropy of the system in terms of the number of gas molecules N and the total number of lattice sites n.

(ii) Show that, when $x = N/n \to 0$ (i.e., low gas density), the entropy change in response to the volume change from V_1 to V_2 is

$$\Delta S = Nk_B \ln(V_2/V_1).$$

(iii) Show that, when $x \to 0$, the lattice model satisfies the ideal-gas law $PV = Nk_BT$.

1.15 Repeat Problem 1.14 by assuming that the gas molecules do not interact with each other and that the m-orientations for each gas molecule on the lattice have different energies $\epsilon_i, i = 1, 2, \dots, m$. How does the molecular orientation influence the relative entropy of the system?

1.16 Consider a lattice model such that each site can be in one of two energy states, 0 or $\epsilon > 0$. Let n be the total number of lattice sites, and N the number of sites with energy ϵ.

(i) How would you define the microstates of this system?

(ii) Show that the lattice sites with different energies follow the Boltzmann distribution

$$\frac{N}{n-N} = \exp\left(-\frac{\epsilon}{k_B T}\right).$$

(iii) Show that the constant-volume heat capacity of the lattice system is given by

$$C_V = \frac{N(n-N)}{n}\frac{\epsilon^2}{k_B T^2}.$$

(iv) Can the absolute temperature be negative? Why?

1.17 Imagine that Maxwell's demon could control the direction of molecules passing through a hole between identical chambers A and B containing a one-component ideal gas. The demon allows all molecules, fast or slow, to pass from B to A but prevent them from passing from A to B. Eventually all molecules will be concentrated in A and a vacuum will be created in B so that a pressure difference is generated without doing any work. Explain whether the process violates the first or the second law of thermodynamics and discuss how this might be implemented with modern technology.

1.18 In a computer, the Memory Address Register (MAR) is responsible for storing either the memory address from which data will be fetched to the Central Processing Unit (CPU), or the address to which data will be sent and stored. Consider a computer operation that reformats a memory register of n bits. Before the operation, the register as a whole could have existed in any of 2^n states. However, after the operation, the register is left in only one state. To maintain the temperature of the computer at this point, how much heat must be released?