1

Introduction

Section 1.1 looks at the similarities and differences between classical thermodynamics and statistical thermodynamics. Then, in Section 1.2, we see several examples of phenomena that are beautifully described by statistical mechanics. Section 1.3 lists practices of notation adopted by this book.

1.1 Classical Thermodynamics and Statistical Thermodynamics

Classical thermodynamics, when applied to a closed system, starts with two fundamental laws. The first law of thermodynamics accounts for a balance of energy:

\[ d'Q + d'W = dU \]  \hspace{1cm} (1.1)

where the system receives heat \( d'Q \) and work \( d'W \) to change its internal energy by \( dU \) (see Figure 1.1). The prime in “d’” indicates that the quantity may not be a thermodynamic variable, i.e. not expressed as a total derivative. When the volume of the system changes from \( V \) to \( V + dV \), \( d'W = -p \, dV \), where \( p \) is the pressure.

The second law of thermodynamics expresses \( d'Q \) by a thermodynamic variable, but only when the change is reversible:

\[ d'Q = T \, dS \]  \hspace{1cm} (1.2)

where \( T \) is the temperature. The second law introduces the entropy \( S \).

In classical thermodynamics, we try to find relationships between macroscopic variables, \( S, T, U, V, \) and \( p \). The equation of state is one of the relationships. We also learned different types of energy, specifically, enthalpy \( H \), Helmholtz free energy \( F \), and Gibbs free energy \( G \). These measures of energy are convenient when we consider equilibria under different constraints. For example, at constant \( T \) and \( V \), it is \( F \) that minimizes when the system is at equilibrium. Certainly, we can always maximize \( S \) of the universe (system + the
surroundings), but knowing the details of the surroundings is not feasible or of our concern. Rather, we want to focus on the system, although it is the maximization of the entropy of the universe that dictates the equilibrium of the system. People have devised $F$ for that purpose. If we minimize $F$ of the system under given $T$ and $V$, we are equivalently maximizing $S$ of the universe. Likewise, $G$ minimizes when the system's temperature and pressure are specified.

As you may recall, classical thermodynamics does not need to assume anything about the composition of the system – whether it is a gas or liquid, what molecules constitute the system, and so on. The system is a continuous medium; and it is uniform at all length scales, if it consists of a single phase. In other words, there are no molecules in this view.

Statistical thermodynamics, in contrast, starts with a molecule-level description of the system – what types of molecules make up the system, whether interactions are present between molecules, and, if they are, how the interaction depends on the distance between molecules, and so on. Furthermore, statistical thermodynamics specifies microscopic states of the molecules, for example, their positions and velocities. If the molecules are monatomic, specifying the positions and velocities may be sufficient for our purposes. When the molecules are diatomic, however, we need to specify the states of rotation and vibration as well. If the molecule is polyatomic, specifying these states becomes more complicated. Even for a system of monatomic molecules, specifying the positions and velocities requires an astronomical number of variables. Typically, the number is close to Avogadro's number. Listing and evaluating all the variables is a daunting task. Fortunately, evaluating thermodynamics variables such as $U$, $F$, and $G$ does not require all the details. It is rather the averages of the microscopic variables that count in evaluating the thermodynamic variables, and that is where statistical thermodynamics comes in.

1.2 Examples of Results Obtained from Statistical Thermodynamics

Here, we take a quick look at some of the results of applying statistical thermodynamics to different systems.

1.2.1 Heat Capacity of Gas of Diatomic Molecules

Figure 1.2 shows how the molar heat capacity $C_V/n$ of a gas consisting of diatomic molecules changes with temperature $T$. There are two characteristic
3

Introduction

Figure 1.2 Molar heat capacity $C_V/n$ of a gas consisting of diatomic molecules, plotted as a function of temperature $T$. At $T$ around $\Theta_{\text{rot}}$, the characteristic temperature of rotation, $C_V/n$ increases from $\frac{3}{2}R$ to $\frac{5}{2}R$; and at around $\Theta_{\text{vib}}$, the characteristic temperature of vibration, further increases to $\frac{7}{2}R$.

Table 1.1 Characteristic temperature of rotation, $\Theta_{\text{rot}}$, and characteristic temperature of vibration, $\Theta_{\text{vib}}$, for some diatomic molecules.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>$\Theta_{\text{rot}}$ (K)</th>
<th>$\Theta_{\text{vib}}$ (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H$_2$</td>
<td>87.6</td>
<td>6331</td>
</tr>
<tr>
<td>N$_2$</td>
<td>2.88</td>
<td>3393</td>
</tr>
<tr>
<td>O$_2$</td>
<td>2.08</td>
<td>2239</td>
</tr>
</tbody>
</table>

temperatures $\Theta_{\text{rot}}$ and $\Theta_{\text{vib}}$ (rotation and vibration). Each diatomic molecule has its own $\Theta_{\text{rot}}$ and $\Theta_{\text{vib}}$, and some of them are listed in Table 1.1.

The molar heat capacity is $\frac{3}{2}R$ at $T \ll \Theta_{\text{rot}}$, where $R$ is a gas constant. We see this range only for H$_2$; for other gases, the boiling point is above $\Theta_{\text{rot}}$. As $T$ increases and surpasses $\Theta_{\text{rot}}$, $C_V/n$ increases to $\frac{5}{2}R$. There is a broad range of temperature that gives a nearly constant value of $C_V/n$ before it increases to $\frac{7}{2}R$ as $T$ exceeds $\Theta_{\text{vib}}$. For most diatomic molecules that are gas at room temperature (RT), $\Theta_{\text{rot}} \ll \text{RT} \ll \Theta_{\text{vib}}$, and that is why a gas of diatomic molecules has $C_V/n = \frac{5}{2}R$.

1.2.2 Heat Capacity of a Solid

Figure 1.3 depicts the molar heat capacity $C_V/n$ of a molecular solid (nonionic), plotted as a function of temperature $T$. At low temperatures, $C_V/n \sim T^3$, and increases to a plateau value of $3R$ as $T$ increases. Vibration in a lattice (crystal) accounts for this heat capacity. Einstein attempted to explain the heat capacity in his 1905 paper [1]. His statistical model correctly predicted $3R$, but not $T^3$. It is Debye who explained the $\sim T^3$ dependence by improving the Einstein model [2].

1.2.3 Blackbody Radiation

Anything with a temperature $T > 0$ radiates. A blackbody is a perfect emitter of the radiation (light) and also a perfect absorber. The radiation has different
wavelength components and is not visible unless the wavelength falls in the visible range, 450–750 nm. When the radiation intensity is plotted as a function of \( T \), the curve peaks at some wavelength \( \lambda_{\text{peak}} \) (see Figure 1.4). With an increasing \( T \), \( \lambda_{\text{peak}} \) moves to a shorter wavelength, and the peak intensity increases. Stars exhibit different colors, and it is due to temperature differences. The radiation from the sun peaks at around 500 nm (blue–green), since its surface temperature is around 5800 K. A red star has a lower temperature, and a white star (\( \lambda_{\text{peak}} \approx 300 \text{ nm} \)) has a higher temperature.

The \( \lambda_{\text{peak}} \) decreases as \( \sim T^{-1} \) as \( T \) increases, which is called Wien’s displacement law, discovered in 1893. The profile of the spectrum has tails at both ends. The long-wavelength tail follows \( \sim \lambda^{-4} \), and short-wavelength tail \( \sim e^{-\text{const.}/\lambda} \). The long-wavelength tail was explained using classical electromagnetism, but it could not explain the short-wavelength tail, or the Wien’s law. Max Planck proposed a photon hypothesis – light consists of energy particles called photons, each carrying energy reciprocally proportional to \( \lambda \) – in 1900 [3]. He succeeded in explaining the whole radiation spectrum.

### 1.2.4 Adsorption

This example is more chemical than are the preceding examples. When a clean surface (glass, graphite, etc.) is exposed to a vapor, some molecules adsorb onto the surface (Figure 1.5a). The surface coverage \( \theta \) (fraction of surface covered with the molecules) increases with an increasing partial pressure \( p \) of the vapor
The adsorption phenomenon can be explained by reaction kinetics, but statistical thermodynamics provides a molecular-level description of the isotherm. For example, we can estimate, from the experimentally obtained isotherm, the cohesive energy of adsorption per molecule.

### 1.2.5 Helix–Coil Transition

A polypeptide is a polymer of identical amino acid residues. For example, poly(l-lysine) is a polymer of L-lysine. The polypeptide adopts a helix conformation, a random-coil conformation, or a mixture of them (a part of the polymer chain is in helix conformation), see Figure 1.6(a). Which conformation the polymer takes depends on the environment such as temperature and solvent. Figure 1.6 (b) is a sketch for a plot of percent helix as a function of temperature. The polypeptide is benzyl ester of poly(glutamic acid), and therefore soluble in a polar organic solvent. At low temperatures, nearly all of the polymer is in a coil conformation, and gradually changes to an all-helix conformation as the temperature rises. It may appear counterintuitive that the...
ordered state of the polypeptide is seen at high temperatures, rather than at low temperatures.

The helix conformation is made possible by hydrogen bonds between a donor (H atom in an amide bond) and an acceptor (O atom in the amide bond) several residues away along the chain. If there is a mechanism that supersedes this intrachain hydrogen bond, then the chain may adopt a coil conformation. The latter can occur if the solvent molecules provide a stronger hydrogen bond to the H atoms and O atoms of the amide bonds. It is a competition between the two types of hydrogen bonds that gives rise to the inverted temperature dependence of the percent helix.

1.2.6 Boltzmann Factor

You would have learned about the Boltzmann factor, \( e^{-\Delta E/(RT)} \), where \( \Delta E \) is the energy difference per mole, without being shown its derivation. The Boltzmann factor appears in many different situations. It appears, for example, in the barometric formula, \( p(h) = p(0)e^{-Mgh/(RT)} \), where \( p(h) \) is the pressure at altitude \( h \), \( M \) is the molar mass of the gas, and \( g \) is the gravitational constant. The Debye–Hückel theory for electrolyte solutions and the Gouy–Chapman theory for colloidal suspensions also use the Boltzmann factor. In nuclear magnetic resonance (NMR), the population of an up spin \( (+ \frac{1}{2}) \) of a magnetic dipole \( \mu \) in magnetic field \( B \) with respect to the down spin \( (- \frac{1}{2}) \) is \( \exp(N_A \mu B/RT) \), where \( N_A \) is the Avogadro’s number.

Statistical thermodynamics derives the Boltzmann factor from fundamental hypotheses. We learn the hypotheses and the derivation in Chapter 4.

1.3 Practices of Notation

This section lists some practices used in this book.

(1) **Symbols:** This book uses italic symbols for variables; roman-typefaced symbols are not variables. For example, the base of the natural logarithm is \( e \), not \( e \). Likewise, the circumference ratio is \( \pi \), not \( \pi \).

(2) **\( O(x^n) \):** represents a quantity proportional to \( x^n \): It may include higher or lower order terms. For example, the Taylor expansion of \( e^x \) is \( 1 + x + \frac{1}{2}x^2 \), up to \( O(x^3) \).

(3) **Limit and asymptote:** We strictly distinguish them. We often consider behaviors of thermodynamic functions at low temperatures and at high temperatures. We evaluate the limiting value of a function \( f(T) \) as \( T \to 0 \), and this is the low-temperature limit. We also evaluate the dominant term of the function when \( T \) is low and when \( T \) is high. Such a term retains the temperature dependence and is called an asymptote.
For example, the molar heat capacity $C_V/n$ of a molecular solid has the low-temperature limit of zero, but its low-temperature asymptote is $C_V/n \sim T^3$. The high-temperature limit of $C_V/n$ is $3R$.

Let us consider the following function $f(x)$ for a practical example.

$$f(x) = \ln(1 + \cosh x), \quad x \geq 0$$

(1.3)

The small-$x$ limit is $\ln 2$. There is no large-$x$ limit, as $f(x) \to \infty$ as $x \to \infty$. However, we can get the large-$x$ asymptote:

$$f(x) \cong x - \ln 2$$

(1.4)

Figure 1.7 shows a plot of $y = f(x)$ and its large-$x$ asymptote. The plot of $y = f(x)$ runs close to the asymptote that is a straight line.

The leading term of the large-$x$ asymptote is $x$, but including the constant makes the asymptote a better approximation. The asymptote, Eq. (1.4), neglects $O(e^{-x})$. 

Figure 1.7 Plot of $y = f(x) = \ln(1 + \cosh x)$ (solid line) and its large-$x$ asymptote (dashed line).