

# 5

---

## Statistical Mechanics

*Perhaps, the atomic hypothesis will be replaced by another some day—perhaps, but not probably.*

Ludwig Boltzmann

In this chapter, the macroscopic world of thermodynamics is linked to the microscopic world of atoms and molecules through the idea of the probability of a macroscopic state. The basic assumption of statistical mechanics is that this probability is proportional to the number of microscopic arrangements that can give the macroscopic state. This number is calculated for distributions in space and among energy levels. The Boltzmann distribution is derived from the second law of thermodynamics. The partition function is used to obtain formulas to calculate thermodynamic functions from molecular properties. The development is applied to a system with only translational energy to give the partition function and thermodynamic properties of the monatomic ideal gas. The Maxwell velocity distribution and several of its averages from kinetic theory are calculated. Extension of the theory to polyatomic molecules in the classical limit gives the classical equipartition theorem. Foreshadowing the discussion of steady-state systems, thermal transpiration is treated by kinetic theory. Statistical theory is

applied to the model of the freely jointed chain to discuss the mechanical properties of macromolecules.

## 5.1 The Microscopic World

Our discussion of thermodynamics has been thoroughly grounded on observations on the real macroscopic world. It is not necessary to be aware of the existence of atoms in order to discuss and use the first, second, and third laws of thermodynamics. The basis for the interpretation of thermodynamics in terms of the microscopic world of atoms, molecules, and ions was first given by Ludwig Boltzmann in 1896. However, at that time, there was so little direct evidence for the existence of these microscopic particles that Boltzmann's ideas were much disparaged and only slowly gained acceptance. The lack of appreciation of his theories may have been a contributing factor in Boltzmann's suicide in 1906. If Boltzmann had not killed himself, he might have lived long enough to see the general acceptance of his concepts and the increased utility of thermodynamics derived from the ability to calculate macroscopic thermodynamic functions from microscopic molecular properties. In this chapter, we will introduce some of the basic ideas behind the relation of thermodynamic properties to microscopic properties. This subject is known as statistical mechanics because it involves averaging over the properties of the microscopic entities in order to obtain the macroscopic properties.

One problem that arises is that although Newtonian mechanics is sufficient to describe the overall translational motion of particles of the size of atoms and molecules, quantum mechanics is required to describe their rotational and internal motion. Quantum mechanics is essential in dealing with the motion of particles as small as electrons. Because knowledge of quantum mechanics is not assumed of the reader of this book, we will be content to develop the framework into which quantum mechanical results can be later be inserted. We will apply this framework to two systems that can be treated classically, namely the monatomic ideal gas and polymer chains.

It is fairly simple to explain the first law of thermodynamics on the basis of atoms. (We will use the term "atom" to generically denote the microscopic particles: atoms, molecules, ions, and electrons.) Newtonian mechanics is known to conserve energy<sup>1</sup> and when we add huge numbers of interactions, each conserving energy, the energy of the sum should also be conserved. The second law, which tells us the direction of real processes, is much more difficult to explain by Newtonian mechanics. The direction of a Newtonian process depends on its initial conditions. If we reverse the velocities of all of the interacting particles, the process runs backward. The reason for this is that Newton's second law involves acceleration, which is a second derivative with

respect to time and thus is unchanged upon reversing the “direction” of time. We shall see that Newtonian mechanics does have a forward direction in a statistical or averaged sense.

## 5.2 The Joule Process

In order to gain a sense of the statistical aspects of thermodynamic processes, let us look in detail at one simple process—the Joule expansion of an ideal gas. This spontaneous process involves the rapid (and thus adiabatic) expansion of a gas into vacuum (thus producing no work). It is described by the Joule coefficient,  $\mu_J \equiv (\partial T / \partial V)_U$ , which we have shown equals  $-(1/C_V)(\partial U / \partial V)_T$  and is zero for an ideal gas. For an ideal gas, energy depends only on temperature and not on volume, and so the Joule process is isothermal as well. Because the Joule expansion is adiabatic and without work, the system is isolated and there are no changes in the surroundings. The driving force for this process must come from some change in the system. In [Chapter 3](#), we called this driving force the entropy increase of the system. From a microscopic point of view, why does the gas spontaneously expand?

We will find it easier to consider the reverse Joule process, which is not spontaneous. We ask the following: If gas occupies an entire container, why at a later time do we *not* find it contracted into some fraction of this container? Choosing a halving of volume, let us look at this process from a microscopic point of view. We will first consider a system containing only four argon atoms ( $N = 4$ ), which initially may be anywhere in the container of volume  $V$ . In the reverse Joule process, the atoms all move to the left half of the container. We could try to explain the nonoccurrence of this process using Newton’s laws, by following the trajectories of the particles. However, this would require a computer and we would find that if there were any uncertainties in the initial conditions, after a short time (say 1 s) we could claim very little accuracy for our calculation.<sup>2</sup> This leads us to give up our attempt for a dynamical description of this process and think about how we might investigate it experimentally.

For an experimental investigation, we could make use of a “magic camera” that could photograph individual argon atoms and take a million photographs of the contents of the box. The camera would be performing a time average of the motion of the atoms in the box.<sup>3</sup> With a million photos and only four atoms in the box, it is likely that we would find some photos with all the atoms on the left side of the box. How would we interpret this result? A reasonable interpretation would be in terms of probability. We could count the pictures that showed all four atoms on the left side of the box (consistent with our definition of the final state of the reverse Joule process) and divide it by the total number of pictures, where the four atoms could be anywhere in the box (consistent with the initial state). This ratio

would be taken as the fraction of the time that the system spent in the final state of the reverse Joule process, and we would assume that as the number of photos became infinite, the ratio would approach the *probability* of the spontaneous occurrence of the final state.

Unfortunately, we do not have a magic camera, and even if we did, the bill for developing the number of photos required would be prohibitive for any macroscopic number of particles. This leads us to attempt to figure out the result of this experiment without doing it. Because we have already given up on the possibility of following the microscopic behavior of the system over time, we are not able to calculate time averages. Instead, we adopt the procedure of calculation using ensemble averages. By an *ensemble*, we mean a large number of systems, each satisfying the macroscopic definition of the initial state of our reverse Joule process (fixed  $N$ ,  $V$ , and  $U$ ). An ensemble in which every system has the same  $N$ ,  $V$ , and  $U$  is appropriate for isolated systems and is known as a *microcanonical ensemble*. Every microscopically distinguishable arrangement of the atoms is included once in the ensemble. We then make the basic postulate of statistical mechanics:

The probability of a macroscopic state is proportional to the number of distinguishable microscopic configurations that are consistent with the definition of that state.

We can justify that the configurations must be distinguishable, because quantum mechanics asserts that there can be no real difference between configurations that cannot be distinguished. Of course, Boltzmann predated quantum mechanics, but he realized that distinguishable configurations were necessary for agreement with observation.

The key word for the calculation in the basic postulate is “number.” What we are claiming is that there is no difference in probability between any distinguishable microscopic configuration of the system. All we have to do is count the possible microscopic configurations corresponding to a state. Therefore, what the microscopic description of thermodynamics comes down to in the end is just counting!

Substituting an ensemble average for the time average rests on a number of assumptions. It is necessary that all configurations represented in the ensemble be accessible during the motion over time of the system. In addition, no distinguishable configuration can be favored over other configurations in the motion over time. It is difficult to see why some configurations would not be accessible in the motion over time and what criteria could be used to favor certain configurations in a system defined only in terms of  $N$ ,  $V$ , and  $U$ . Although there has been considerable discussion of these assumptions in the literature, we will just accept the above basic postulate.<sup>4</sup>

In counting distinct configurations, we must realize that coordinate space is continuous, and because positions that differ infinitesimally can be distinguished, theoretically we can enumerate an infinite number of configurations. Ratios of infinities are not much use to us for calculating probabilities! The way to get around this is to divide the container into a large number of microscopic compartments, say  $M$  for each half of the box. The size of the compartments should be large compared to that of an atom but small enough that there are many more compartments than atoms and little probability that two atoms are in the same compartment.<sup>5</sup> This will make our counting much easier. A configuration is then defined by which of the  $M$  compartments is occupied (by the center of the atom). Note that even though the atoms are indistinguishable, these configurations would give distinguishable photographs with our magic camera.

For the reverse Joule process, initially the atoms can occupy  $2M$  compartments. Counting configurations for four atoms in the initial system is equivalent to asking how many ways can four indistinguishable atoms occupy  $2M$  compartments, with no more than one atom in a compartment? This is a fundamental problem in probability theory. The way it is solved is by first taking all possible permutations (orderings) of the  $2M$  compartments [there are  $(2M!) = 2M(2M - 1)(2M - 2) \cdots 1$  of these] and then dividing by the possible permutations of the filled  $(4!)$  and empty  $[(2M - 4)!]$  compartments, both of which are indistinguishable sets. We will call the number of configurations consistent with a macroscopic state  $\Omega$ . We thus have for the initial state

$$\Omega_i = \frac{(2M)!}{4!(2M - 4)!} = \frac{2M(2M - 1)(2M - 2)(2M - 3)}{4!} \approx \frac{(2M)^4}{4!} \quad (1)$$

where the final approximation results from  $M$  being much larger than 4. Likewise, the number of configurations consistent with the final state (four atoms in  $M$  compartments) is

$$\Omega_f = \frac{(M)!}{4!(M - 4)!} \approx \frac{M^4}{4!} \quad (2)$$

This gives

$$\frac{\Omega_f}{\Omega_i} = \frac{1}{2^4} = \frac{1}{16} \quad (3)$$

as the ratio of the number of configurations of the final state to that of the initial state. Now let us say that instead of four atoms in the box, there are  $N$  atoms. Redoing our calculation, we would get

$$\frac{\Omega_f}{\Omega_i} = \frac{1}{2^N} \quad (4)$$

With  $N$  corresponding to any macroscopic amount of matter (say  $1 \text{ nmol} = 6 \times 10^{14}$  atoms), this is such a small number that the probability of all of the atoms being in the left-hand side of the box is totally negligible. This is a dramatic example of the microscopic meaning of the second law. Although there is a *possibility*, in a macroscopic amount of a gas, of all the atoms suddenly finding themselves in one-half of a container, the *probability* of this happening is so small that we can completely neglect it. The reason that the reverse Joule process does not spontaneously occur is therefore a statistical one. It is very, very unlikely that it would ever occur. Conversely, the Joule process, the spontaneous expansion of a gas, is statistically very likely.

For systems containing macroscopic numbers of particles, the probability of finding even 51% of the atoms (or any measurable difference) in one-half of the box is also negligible. We can state the following:

For macroscopic numbers of molecules, states whose properties differ measurably from the state with the maximum number of configurations have negligible probability. The properties of the system are therefore just the properties of the state with the maximum number of configurations.

Note that from our calculation on the four-atom case, we can see that for a small number of particles, the probability of appreciable percentage fluctuations from the uniform distribution is more probable.

Because both entropy and the number of configurations indicate the direction of spontaneous processes, these must be related. What is the relationship between these quantities? Although both  $\Omega$  and  $S$  tend to maximum values in real processes,  $S$  cannot be equal (or proportional) to  $\Omega$  because  $\Omega$  is not an extensive function. For example, if in a given state of the container considered here, there are  $\Omega_l$  configurations for the left side of the container and  $\Omega_r$  configurations for the right side of the container, then there are  $\Omega_l \Omega_r$  configurations for the entire container, because each left-side configuration can go with any right-side configuration. In order to convert this product into the sum needed for an extensive property, we set  $S$  proportional to  $\ln \Omega$ ,<sup>6</sup> giving Boltzmann's famous relationship<sup>7</sup>

$$S = k \ln \Omega \quad (5)$$

To find  $k$ , the Boltzmann constant, we calculate  $\Delta S$  for the halving of the volume of an ideal gas, considered here in the reverse Joule process. From Eq. (4), if  $\Omega_i = C$ , then  $\Omega_f = 2^{-N} C$ , and

$$\Delta S = k \ln (2^{-N} C) - k \ln C = k \ln 2^{-N} = -Nk \ln 2 \quad (6)$$

However, we have seen in [Chapter 3](#) that the entropy change for the isothermal (and constant energy) halving of a volume of  $n$  moles of an ideal gas is  $-nR \ln 2$ . If our statistical result is to agree with our thermodynamic result,

then  $Nk = nR$  or  $k = R(N/n)^{-1} = R/N_A$ . This is the same constant that arose in calculating the average energy of a molecule in the Bernoulli model of the ideal gas in Section 1.4 of [Chapter 1](#). In fact, the Boltzmann constant will arise whenever we calculate macroscopic properties from a microscopic model of a system. The value of Boltzmann's constant is  $1.38 \times 10^{-23}$  J/K.

We have discussed Eq. (5) in connection with a system at constant  $N$ ,  $V$ , and  $U$  (a microcanonical ensemble). There is no reason, however, why the same relationship between  $\Omega$  and  $S$  should not hold under other constraints—in particular, at constant  $N$ ,  $V$ , and  $T$  (called a *canonical ensemble*).

Equation (5) also gives a microscopic explanation of the third law of thermodynamics. In most perfectly crystalline materials, there is a unique arrangement that has the lowest energy. As the temperature is lowered and energy is removed from the material, it tends to this lowest-energy state. With a unique lowest-energy state,  $\Omega = 1$ , and  $S = 0$  is approached as  $T \rightarrow 0$ . Some species (such as CO, discussed in [Chapter 4](#)) have residual entropy, due to an alternative arrangement that is very close in energy to the lowest-energy configuration. In this case, it may be that at temperatures at which the lowest energy configuration becomes favored by the Boltzmann distribution, there is no longer sufficient thermal energy to overcome the kinetic activation barrier to make a transition to the lowest energy state. The random distribution between the two low-energy arrangements then becomes “frozen in” as  $T \rightarrow 0$ .

**Example 1.** Calculate the residual entropy at  $T = 0$  for carbon monoxide.

*Solution:* Carbon monoxide has a small electric dipole moment (approx 0.1 Debye), which gives the molecules an energetically preferred orientation as  $T \rightarrow 0$ . However, this dipole moment is so small that the preference is not appreciable until very low temperatures, and the random orientation of the molecules (the dipole has equal probability of pointing in one direction or its opposite) remains as the temperature is lowered. For a mole of CO, each molecule can point in either of two directions and there are  $2^{N_A}$  configurations that are about equally probable. This model predicts a residual entropy of

$$S(0) = k \ln 2^{N_A} = R \ln 2 = 5.76 \text{ J/mol K}$$

which is close to the measured value of 5.0 J/mol K for CO.

Note that in solid CO, a configuration is determined by *which* CO molecules have the less favorable orientation, not by just how many molecules have this orientation. This is because, unlike in a gas, in a solid, molecules can be identified by their position in the lattice. Thus, configurations in which the same number, but different, molecules are less favorably oriented can be distinguished.

Using Eq. (5), we can write

$$\frac{\Omega_i}{\Omega_{\max}} = \exp\left(\frac{S_i - S_{\max}}{k}\right) = \frac{P_i}{P_{\max}} \quad (7)$$

where “max” refers to the state with the maximum number of configurations (i.e., the equilibrium state). Because numbers of configurations are proportional to the probability of states, Eq. (7) also gives the probability of a spontaneous fluctuation of the system away from the equilibrium state with maximum entropy to a state of lower entropy. Entropy differences for macroscopic processes are usually of the order of  $R$ . Thus the probability of a spontaneous fluctuation of a system by a macroscopically measurable amount from the equilibrium state is of the order of  $\exp(-N_A)$ , a very, very small number.

### 5.3 Distribution Among Energy States

So far, we have considered how molecules are distributed among compartments of equal volume. These compartments can be thought of as molecular states of equal energy, and we postulated that there is equal probability of occupying such states. How do molecules occupy states when they are not at equal energy? We will consider a system comprised of noninteracting particles in contact with a heat reservoir at temperature  $T$ . Because the particles are noninteracting, each particle has energy levels that are independent of the other particles. Energy being continuous in classical mechanics,<sup>8</sup> we will envision small ranges of energy, which we designate as  $\epsilon_j$ . We seek the *occupation numbers*,  $N_j$  of these ranges, such that the sum of the  $N_j$  equals the total number of molecules in the system,  $N = \sum_j N_j$ . We will investigate a process in which we change the occupation numbers.

We will use the microscopic definition of entropy, Eq. (5), to calculate the entropy change of the system in this process. In order to do this, we must calculate the number of distinguishable microscopic configurations corresponding to the initial and final distributions of the molecules over their energy levels. Although the molecules are indistinguishable, we will assume that they can be numbered, perhaps by their position in the container. This way of calculating the number of configurations is called *Boltzmann statistics*. With  $N$  numbered molecules, the number of ways of assigning a set of occupation numbers  $\{N_0, N_1, \dots, N_j, \dots\}$  to the molecules is

$$\Omega = \frac{N!}{N_0!N_1!\dots N_i!\dots} = \frac{N!}{\prod_i N_i!} \quad (8)$$



The correctness of this very important formula is illustrated for the example  $N = 3$ ,  $N_0 = 2$ ,  $N_1 = 1$ . There are obviously three configurations for this distribution, corresponding to the three choices of which molecule we place in state 1. The general formula, Eq. (8), is arrived at by including a partition in the ordered listing of the molecules as follows: 1 2|3. All  $3!$  permutations of the three molecules will now place different molecules in different positions, but these will include configurations such as 2 1|3, which differs from 1 2|3 only by an indistinguishable permutation of identical molecules in the same state. Therefore, we must divide  $3!$  by the  $2!$  possible permutations of the identical molecules in state 0 to get our final result.

Quantum mechanics tells us that particles in the gas phase cannot be distinguished by their position; thus, the above-considered example cannot correspond to more than a single configuration. In fact, in quantum systems, there are two different ways of counting the number of configurations, each appropriate for different types of particles.<sup>9</sup> Gibbs first suggested that for indistinguishable particles, the result of Eq. (8) be divided by  $N!$  to give *corrected Boltzmann statistics*. Although this procedure cannot be justified in any rigorous manner, it can be shown to approach the correct results for both types of quantum particles as the temperature is increased. Dividing the result of Eq. (8) by  $N!$  is seen to have no effect on the results derived in this section.

From Eq. (8), the minimum number of configurations is one, which occurs when all the molecules are in the same state; the maximum number of configurations is  $N!$ , which occurs when each  $N_i$  is either 0 or 1 (because  $1! = 0! = 1$ ). In other words, a broad distribution of molecules over energy states has a higher entropy than a narrow distribution. Why then, in order to satisfy the second law, does the system not achieve a state where there is no more than a single molecule in each state? The reason for this is that such a distribution would require that the system have a large amount of energy, so that high-energy levels could be occupied. For an isolated system, the energy is limited and the system will achieve the broadest distribution consistent with that energy. For a closed system in contact with a thermal reservoir at temperature  $T$ , the system can increase the breadth of its distribution and its entropy by extracting energy from the reservoir. However, this will lower the entropy of the surroundings, and because  $\Delta S_{\text{sur}} = -q/T$ ,  $\Delta S_{\text{sur}}$  will be lowered further, the lower the temperature of the reservoir and the more energy that is transferred to the system.

To find the equilibrium distribution, we consider a system of fixed volume<sup>10</sup> and number of particles in contact with a thermal reservoir at temperature  $T$ . By the second law, at equilibrium there must be a zero entropy change of the universe for all infinitesimal changes in the system, because  $S_{\text{univ}}$  is a maximum at equilibrium. In particular, for the change involved in increasing the number of molecules in the  $j$ th state (with  $\varepsilon = \varepsilon_j$ ) of the system by 1, while decreasing the number of molecules in the lowest state of the system (with  $\varepsilon = 0$ )

by 1, thermal energy of the amount  $\varepsilon_j$  must be transferred to the system. This produces an entropy decrease in the surroundings of

$$\Delta S_{\text{sur}} = -\frac{\varepsilon_j}{T} \quad (9)$$

Because the system volume is constant, no energy can be transferred to the system by compression work; we assume that no other work is done on the system.) Using Eqs. (5) and (8), the entropy change of the system becomes

$$\Delta S = S_f - S_i = k \ln \Omega_f - k \ln \Omega_i = k \ln \left( \frac{\prod_{j,i} N_{j,i}!}{\prod_{j,f} N_{j,f}!} \right) \quad (10)$$

Note that for the process that we are considering, one molecule going from the lowest ( $\varepsilon = 0$ ) state to the  $j$ th state, all terms in the numerator and denominator cancel except the  $N_0!$  and  $N_j!$  terms.  $\Delta S$  then becomes

$$\Delta S = k \ln \left( \frac{N_{0,i}! N_{j,i}!}{(N_{0,i} - 1)! (N_{j,i} + 1)!} \right) = k \ln \left( \frac{N_{0,i}}{(N_{j,i} + 1)} \right) \approx k \ln \left( \frac{N_0}{N_j} \right) \quad (11)$$

In the last step, we have assumed that the range of energy chosen for each level is large enough that the  $N$ 's are very large numbers, giving  $N_{j,i} + 1 = N_{j,f} \approx N_{j,i} \equiv N_j$ .

If the system is at equilibrium, the entropy change (of the universe) for this process must be zero:

$$\Delta S_{\text{univ}} = 0 = -\frac{\varepsilon_j}{T} + k \ln \left( \frac{N_0}{N_j} \right) \quad (12)$$

or

$$\frac{N_j}{N_0} = \exp \left( -\frac{\varepsilon_j}{kT} \right) \quad (13)$$

When using this distribution function, we usually know  $N$ , the total number of molecules, rather than  $N_0$ .  $N$  is just the sum over molecules occupying all states

$$N = \sum_{\text{states}} N_j = N_0 \sum_{\text{states}} \exp \left( -\frac{\varepsilon_j}{kT} \right) = N_0 q \quad (14)$$

where  $q$  is known as the *molecular partition function*:

$$q \equiv \sum_{\text{states}} \exp \left( -\frac{\varepsilon_j}{kT} \right) \quad (15)$$

The result gives the *Boltzmann distribution*:

$$f_j = \frac{N_j}{N} = \frac{1}{q} \exp\left(-\frac{\varepsilon_j}{kT}\right) \quad (16)$$

where  $f_j$  is the fraction of the molecules that are in state  $j$ .

The Boltzmann distribution is *normalized*, in that the sum of  $f_j$ , the fraction of the molecules that are in state  $j$ , over all states is unity:

$$\sum_{\text{states}} f_j = 1 \quad (17)$$

There is an alternative way of writing the Boltzmann distribution. Recognizing that there may be two or more states with the same energy (e.g., the six  $2p$  states or ten  $3d$  states of an atom), we introduce the quantity  $g_j$ , the *degeneracy* of states at energy  $\varepsilon_j$ . To find the fraction of molecules that have energy  $\varepsilon_j$ , we just multiply the fraction of molecules that are in one state at the energy [Eq. (16)] by the number of states,  $g_j$ , at the energy

$$f_j = \frac{g_j}{q} \exp\left(-\frac{\varepsilon_j}{kT}\right) \quad (18)$$

If the states span a continuous, rather than a discrete, energy range, we can introduce a continuous *distribution function*,  $f(\varepsilon) d\varepsilon$ , the fraction of molecules that have energy in a range  $d\varepsilon$  around  $\varepsilon$ :

$$f(\varepsilon) d\varepsilon = \frac{g(\varepsilon)}{q} \exp\left(-\frac{\varepsilon}{kT}\right) d\varepsilon \quad (19)$$

where  $g(\varepsilon)$  is the number of states per unit energy at  $\varepsilon$  and

$$q = \int_0^\infty g(\varepsilon) \exp\left(-\frac{\varepsilon}{kT}\right) d\varepsilon \quad (20)$$

In this case, normalization becomes

$$\int_0^\infty f(\varepsilon) d\varepsilon = 1 \quad (21)$$

A variable (or variables) other than energy can be used to represent a continuous range of states. Calling this or these variables  $h$ , we have

$$f(h) dh = \frac{g(h)}{q} \exp\left(-\frac{\varepsilon(h)}{kT}\right) dh \quad (22)$$

and

$$q = \int g(h) \exp\left(-\frac{\varepsilon(h)}{kT}\right) dh \quad (23)$$

One problem that arises in dealing with continuous distributions is that because states are infinitesimally close together, the number of states in any finite energy range is infinite. Such a distribution cannot be normalized unless an arbitrary decision is made on what range of energy (or other variables) corresponds to the definition of a “state.” We will return to this question later in this chapter.

From Eq. (13), the Boltzmann factor,  $\exp(-\varepsilon_j/kT)$ , is the probability of a state at energy  $\varepsilon_j$  being occupied compared to the probability of a zero-energy state being occupied. At  $T = 0$ , the only states that can be occupied are the states at zero energy. At higher temperatures, we must add to these  $g_0$  states the sum of the  $g_j$  states at energy  $\varepsilon_j$ , weighted by the probability of states at this energy being occupied (i.e., the Boltzmann factor) for all energy levels. Performing this addition is just a calculation of the partition function. It follows that the partition function is a measure of the number of states thermally available at temperature  $T$ .

In addition to being a function of  $T$ , the partition function is also a function of  $V$ , on which the quantum description of matter tells us that the molecular energy levels,  $\varepsilon_j$ , depend. Because, for single-component systems, all intensive state variables can be written as functions of two state variables, we can think of  $q(T, V)$  as a state function of the system. The partition function can be used as one of the independent variables to describe a single-component system, and with one other state function, such as  $T$ , it will completely define the system. All other properties of the system (in particular, the thermodynamic functions  $U, H, S, A$ , and  $G$ ) can then be obtained from  $q$  and one other state function.

## 5.4 Thermodynamic Functions from the Partition Function

Unlike the case for entropy, there is no absolute zero for energy. We measure energy with respect to a chosen value,  $U_0$ , per mole of each substance at 0 K (corresponding to each molecule in its lowest-energy state), to which we add the excitation energy of the molecules:

$$U_m = U_0 + N_A \langle \varepsilon \rangle = U_0 + N_A \sum_{\text{states}} f_j \varepsilon_j = U_0 + \frac{N_A}{q} \sum_{\text{states}} \varepsilon_j e^{-\varepsilon_j/kT} \quad (24)$$

where  $\langle \varepsilon \rangle$  is the average energy per molecule. Noting that

$$\begin{aligned} \left( \frac{\partial \ln q}{\partial T} \right)_V &= \frac{1}{q} \left( \frac{\partial q}{\partial T} \right)_V = \frac{1}{q} \left( \frac{\partial \sum_{\text{states}} e^{-\varepsilon_j/kT}}{\partial T} \right)_V \\ &= \frac{1}{qkT^2} \sum_{\text{states}} \varepsilon_j e^{-\varepsilon_j/kT} \end{aligned} \quad (25)$$

we get, with  $kN_A = R$ ,

$$U_m = U_0 + RT^2 \left( \frac{\partial \ln q}{\partial T} \right)_V \quad (26)$$

We hold  $V$  constant in these derivatives because quantum mechanics tells us that the molecular energy levels are functions of  $V$ .

In order to calculate the entropy, we use its fundamental definition [Eq. (5)]. For distinguishable particles (Boltzmann statistics), we use formula (8) for the number of configurations:

$$S_m^{\text{dis}} \equiv k \ln N_A! - \sum_i k \ln N_i! \quad (27)$$

with  $\sum_i N_i = N_A$ . *Stirling's approximation*, for the logarithm of the factorial of a large number,<sup>11</sup> is

$$\ln A! = A \ln A - A \quad (28)$$

giving

$$\begin{aligned} S_m^{\text{dis}} &\equiv kN_A \ln N_A - kN_A - k \sum_i N_i \ln N_i + k \sum_i N_i \\ &= kN_A \ln N_A - k \sum_i N_i \ln N_i \end{aligned} \quad (29)$$

From the Boltzmann distribution function, for  $N = N_A$

$$\ln N_j = \ln N_A - \ln q - \frac{\epsilon_j}{kT} \quad (30)$$

Substituting into Eq. (29),

$$S_m^{\text{dis}} \equiv kN_A \ln N_A - k \sum_i N_i \ln N_A + k \sum_i N_i \ln q + \frac{1}{T} \sum_i N_i \epsilon_i \quad (31)$$

The first two terms in this equation cancel, and the last term is just  $(U_m - U_0)/T$  per mole, yielding

$$S_m^{\text{dis}} \equiv R \ln q + \frac{U_m - U_0}{T} \quad (32)$$

Remembering that we count configurations for indistinguishable particles by dividing the number for distinguishable particles by  $N!$ , we get

$$\begin{aligned} S_m^{\text{indis}} &\equiv R \ln q + \frac{U_m - U_0}{T} - k \ln N_A! = R \ln q + \frac{U_m - U_0}{T} \\ &\quad - k(N_A \ln N_A - N_A) \end{aligned} \quad (33)$$

$$S_m^{\text{indis}} \equiv R \ln \left( \frac{q}{N_A} \right) + R + \frac{U_m - U_0}{T} \quad (34)$$

With these formulas for  $U_m$  and  $S_m$ , equations can be directly derived for the other thermodynamic functions:  $H_m$ ,  $A_m$ , and  $G_m$ . For condensed phases  $H_m = U_m$  to a very good approximation. For gases, assumed ideal,  $H_0 = U_0$  and  $H_m = U_m + RT$ , giving

$$H_m \stackrel{\text{i.g.}}{=} U_0 + RT^2 \left( \frac{\partial \ln q}{\partial T} \right)_V + RT \quad (35)$$

For the Helmholtz free energy,  $A_m = U_m - TS_m$ , giving

$$A_m \stackrel{\text{dis}}{=} U_0 - RT \ln q \quad (36)$$

and

$$A_m \stackrel{\text{indis}}{=} U_0 - RT \ln \left( \frac{q}{N_A} \right) - RT \quad (37)$$

For condensed phases  $G_m \approx A_m$ ; whereas for an ideal gas (indistinguishable),  $G_m = A_m + RT$ ; thus,

$$G_m \stackrel{\text{i.g.}}{=} \mu = U_0 - RT \ln \left( \frac{q}{N_A} \right) \quad (38)$$

In calculations of thermodynamic functions,  $q$  is calculated at the appropriate temperature and volume. For example, for standard thermodynamic functions at 298 K,  $q$  is calculated at 298 K and 24.7 L.

## 5.5 System Partition Functions

We have derived a formula for the molecular partition function by considering a system containing many molecules at equilibrium with a heat bath. We can generalize our statistical mechanics by a gedanken experiment of considering a large number of identical *systems*, each with volume  $V$  and number of particles  $N$  at equilibrium with the heat bath at temperature  $T$ . Such a “supersystem” is called a *canonical ensemble*. Our derivation is the same; the fraction of systems that are in a state with energy  $E_i$  is

$$f_i = \frac{1}{Q} \exp \left( -\frac{E_i}{kT} \right) \quad (39)$$

where  $Q$  is the *system partition function*, given by

$$Q = \sum_{\text{states}} \exp \left( -\frac{E_i}{kT} \right) \quad (40)$$

and the sum is over system states. Alternatively, we can sum over energies rather than states:

$$Q = \sum_{\text{energies}} g(E_i) \exp\left(-\frac{E_i}{kT}\right) \quad (41)$$

where  $g(E_i)$  is the degeneracy, the number of system states at energy  $E_i$ . Usually, for systems, the number of states is so high that they can be considered continuous, and Eq. (41) is written as

$$Q = \int_0^\infty g(E) \exp\left(-\frac{E}{kT}\right) dE \quad (42)$$

where  $g(E)$  is the state density (the number of states between  $E$  and  $E + dE$ ).

Because for systems, unlike particles, there are no requirements for them to be indistinguishable, we use the thermodynamic formulas analogous to those for distinguishable particles. For systems containing  $N$  particles and having volume  $V$ ,<sup>12</sup>

$$U = kT^2 \left( \frac{\partial \ln Q}{\partial T} \right)_{V,N} \quad (43)$$

$$S = \frac{U}{T} + k \ln Q \quad (44)$$

$$A = -kT \ln Q \quad (45)$$

$$\mu = \left( \frac{\partial A}{\partial n} \right)_{T,V} = N_A \left( \frac{\partial A}{\partial N} \right)_{T,V} = -RT \left( \frac{\partial \ln Q}{\partial N} \right)_{T,V} \quad (46)$$

It is difficult to evaluate  $Q$  in cases in which there are interactions between the particles of the system, because, in this case,  $E_i$  is truly an energy level of the entire macroscopic system. In the absence of such interactions,  $E_i$  is just a sum over the energies of the individual particles (atoms, molecules, etc.) comprising the system:

$$E_i = \sum_{k=1}^N \varepsilon_{ik} \quad (47)$$

where  $\varepsilon_{ik}$  is the energy of the  $k$ th particle when the system is in the  $i$ th state. If the particles are distinguishable, each particle can be in any of its states, independent of the other particles, and

$$Q^{\text{dis}} = q_1 q_2 \cdots q_N \quad (48)$$

(Problem 12). If the noninteracting particles comprising the system are identical, then they have the same molecular partition function and

$$Q \stackrel{\text{dis}}{\equiv} q^N \quad (49)$$

If the particles are indistinguishable and identical, this formula can just be divided by  $N!$ , as long as there are many more particle states than particles in the system.

## 5.6 Velocity Distributions

The *kinetic theory of gases* deals with the translational motions of gaseous molecules. Translational motion of molecules can be considered separately from rotational and internal motions and has energy

$$\epsilon_{\text{tr}} = \frac{1}{2}mc^2 = \frac{1}{2}m(v_x^2 + v_y^2 + v_z^2) \quad (50)$$

where  $c$  is the molecular speed. Because classical mechanics is appropriate for treating translational energy in macroscopic containers, all values of velocity are possible and we should use Eq. (19), the distribution function for continuous-energy levels.

$$\begin{aligned} f(\epsilon)d\epsilon &= \frac{g(\epsilon)}{q} \exp\left(-\frac{\epsilon}{kT}\right)d\epsilon \\ &= \frac{g(\epsilon)}{q} \exp\left(-\frac{mc^2}{2kT}\right)d\epsilon \\ &= \frac{g(\epsilon)}{q} \exp\left(-\frac{mv_x^2}{2kT}\right) \exp\left(-\frac{mv_y^2}{2kT}\right) \exp\left(-\frac{mv_z^2}{2kT}\right)d\epsilon \end{aligned} \quad (51)$$

The exponential terms tell us how the molecules are distributed in velocity space, and the degeneracy,  $g(\epsilon) d\epsilon$ , is proportional to the volume in velocity space that gives rise to energies between  $\epsilon$  and  $\epsilon + d\epsilon$ :

$$g(\epsilon) d\epsilon = C 4\pi c^2 dc \quad (52)$$

$C$  is a normalization factor determined by integrating over all of velocity space:

$$\frac{C}{q} 4\pi \int_0^\infty c^2 \exp\left(-\frac{mc^2}{2kT}\right) dc = 1 \quad (53)$$

Two formulas that may be found in most tables of integrals and are very useful in kinetic theory are

$$\int_0^\infty x^{2n} \exp(-\beta x^2) dx = \frac{1}{2} \sqrt{\pi} \frac{(2n)! \beta^{-(n+1/2)}}{2^{2n} n!} \quad (54)$$

and

$$\int_0^\infty x^{2n+1} \exp(-\beta x^2) dx = \frac{1}{2} n! \beta^{-(n+1)} \quad (55)$$



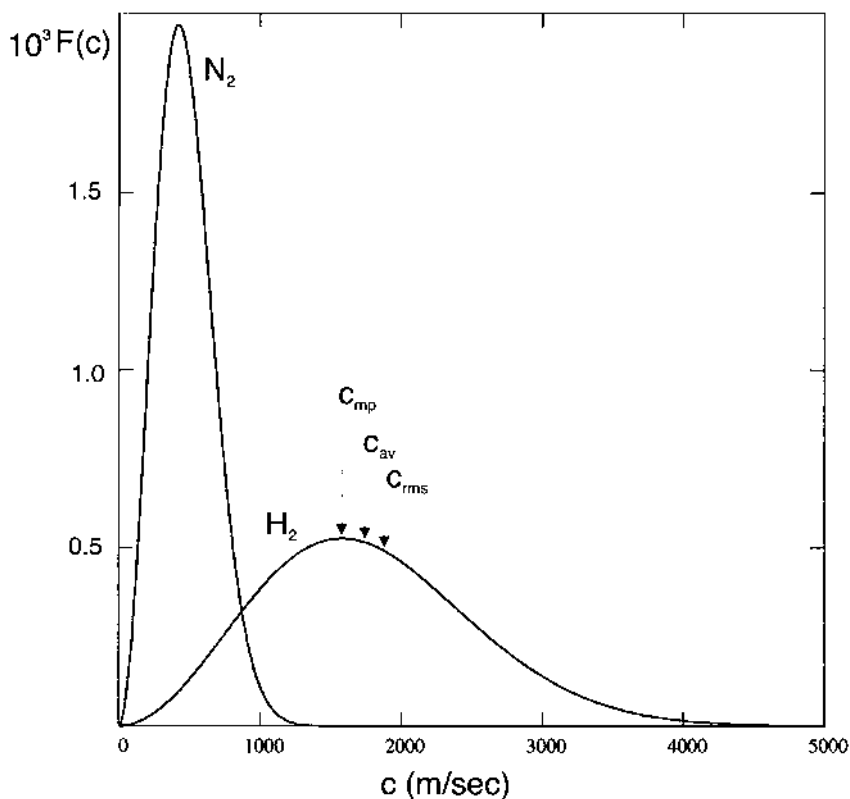
where  $n$  is an integer. From Eq. (53), this gives

$$\frac{C}{q} = \left( \frac{m}{2\pi kT} \right)^{3/2} \quad (56)$$

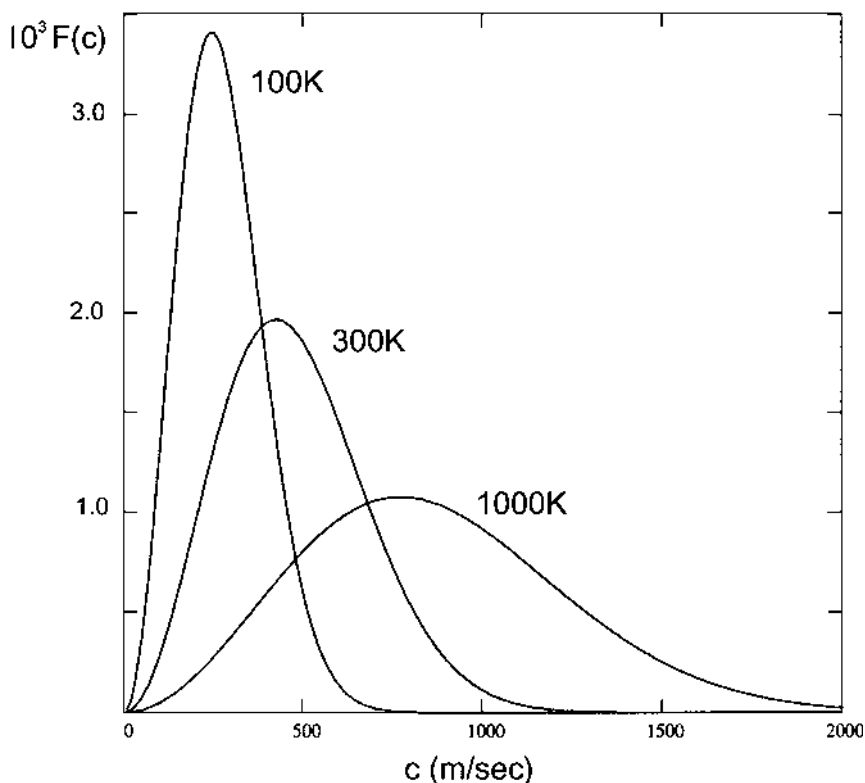
and

$$f(\epsilon) d\epsilon = 4\pi \left( \frac{m}{2\pi kT} \right)^{3/2} c^2 \exp\left(-\frac{mc^2}{2kT}\right) dc = F(c) dc \quad (57)$$

$F(c)$ , the distribution over molecular speeds, is known as the *Maxwell–Boltzmann velocity distribution*. In Fig. 1a, the distribution are shown for  $N_2$  and  $H_2$ , both at 300 K. Note that because the distributions are normalized, the areas under both curves are the same. The most probable speed, average speed, and root-mean-



**Figure 1** (a) Speed distribution of  $N_2$  and  $H_2$  at 300 K.



**Figure 1** (b), Speed distribution of  $N_2$  at 100 K, 300 K, and 1000 K.

square speed are marked on the  $H_2$  distribution curve. In Fig. 1b, distributions are shown for  $N_2$  at 100 K, 300 K, and 1000 K.

Equation (57) is the product of three *one-dimensional velocity distributions*,

$$f(v_i) = \left(\frac{m}{2\pi kT}\right)^{1/2} \exp\left(-\frac{mv_i^2}{2kT}\right) dv_i \quad (58)$$

(where  $i = x, y$ , or  $z$ ), integrated over the spherical shell in velocity space, which encompasses all speeds between  $c$  and  $c + dc$ .

The velocity distribution functions allow us to evaluate the average value of any function of  $v_i$  or  $c$  by integrating over the appropriate distribution. For example, the average molecular speed,  $\langle c \rangle$ , is, using Eqs. (55) and (57),

$$\langle c \rangle = \int_0^\infty c 4\pi \left(\frac{m}{2\pi kT}\right)^{3/2} c^2 \exp\left(-\frac{mc^2}{2kT}\right) dc = \sqrt{\frac{8kT}{\pi m}} \quad (59)$$

Another useful quantity is the rate at which molecules hit a unit area of a wall (or of a hole). Placing the  $x$  axis perpendicular to the wall,  $(N/V)v_x$  molecules with the  $x$ -component of velocity  $v_x$  will hit a unit area of the wall in unit time (all of the molecules in the rectangular parallelepiped of height  $v_x$  and unit cross-sectional area). This quantity has to be averaged over the distribution of  $v_x$  for positive values of  $v_x$  (only molecules moving in the positive direction will hit the wall):

$$\text{Rate}_{\text{wall}} = \int_0^{\infty} \left(\frac{N}{V}\right) v_x \left(\frac{m}{2\pi kT}\right)^{1/2} \exp\left(-\frac{mv_x^2}{2kT}\right) dv_x = \left(\frac{N}{V}\right) \frac{\langle c \rangle}{4} \quad (60)$$

The  $x$  component of kinetic energy transported by these molecules is

$$\text{KE}_x = \int_0^{\infty} \left(\frac{N}{V}\right) v_x \left(\frac{mv_x^2}{2}\right) \left(\frac{m}{2\pi kT}\right)^{1/2} \exp\left(-\frac{mv_x^2}{2kT}\right) dv_x = \left(\frac{N}{V}\right) \frac{\langle c \rangle}{4} kT \quad (61)$$

[Equation (55) is used to evaluate both of these integrals.] Equation (61) means that  $kT$  of  $x$ -component of kinetic energy is transported by each molecule passing through the hole, twice the  $x$  component of kinetic energy that the molecule has in the bulk gas. Because the  $y$  and  $z$  components of kinetic energy that the molecules transport are just the values that they have in the bulk, the average total kinetic energy transported is

$$\langle \text{KE}_{\text{trans}} \rangle = 2kT \quad (62)$$

We will make use of this result in [Chapter 12](#).

## 5.7 A Steady-State Example

Consider two chambers separated by a thin wall in which there is a small hole. We will assume that effusive flow holds for molecules passing through the hole in both directions (i.e., that molecules hitting the hole pass through the wall without collisions with other molecules or with the wall). At thermal equilibrium, both chambers are at the same temperature and, therefore, by Eq. (59), the average speeds of the molecules in the two containers are equal. For mechanical equilibrium, the rate of flow in both directions through the hole must be equal. As a result, from Eq. (60), the density and the pressure in both chambers are equal.

What happens if, by some external means, we maintain the two chambers at different temperatures, say  $T_1$  and  $T_2$ ? In order to do this, we must continually inject heat into the higher-temperature chamber and remove it from the lower-temperature chamber, to counterbalance the tendency of the temperature of the chambers to equalize.<sup>13</sup> In this case as well, equality of flow in both directions through the hole will quickly be attained. The system will then be unchanging in

time, but because the unchanging condition relies upon interactions with the surroundings (the heat transfers), the system is at steady state, rather than equilibrium.

Using Eq. (60) to equate the flows in the two directions at steady state,

$$\left(\frac{N}{V}\right)_1 \langle c_1 \rangle = \left(\frac{N}{V}\right)_2 \langle c_2 \rangle \quad (63)$$

$$\left(\frac{N}{V}\right)_1 \left(\frac{N}{V}\right)_2^{-1} = \frac{P_1 T_2}{P_2 T_1} = \frac{\langle c_2 \rangle}{\langle c_1 \rangle} = \sqrt{\frac{T_2}{T_1}} \quad (64)$$

or

$$\frac{P_1}{P_2} = \sqrt{\frac{T_1}{T_2}} \quad (65)$$

We see that establishing a temperature difference between the chambers has resulted in a density and pressure difference! The movement of material due to a temperature differential is known as *thermal diffusion*, and in the special case where the flow is molecular, it is known as *thermal transpiration*.

## 5.8 Thermodynamic Functions of the Monatomic Ideal Gas

Excitation of the internal energy of monatomic ideal gases requires so much energy that it can be neglected at ordinary laboratory temperatures. The only type of energy that must be considered for such species is that of their translational motion, which we can write as

$$\epsilon_{tr} = \frac{1}{2}mv^2 = \frac{p^2}{2m} = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m} \quad (66)$$

where  $\mathbf{p} = m\mathbf{v}$  is the momentum, which, like velocity, is a vector quantity.

Because the motion of a classical point particle is completely determined by specifying its initial position and momentum, we choose these variables to represent the state of the particle. The six-dimensional space of coordinates and momenta, in which the state of each particle is represented by a point, is called *phase space*. Equation (23) then becomes

$$q = \int g(\mathbf{r}, \mathbf{p}) \exp\left(-\frac{p^2}{2mkT}\right) d\mathbf{r} d\mathbf{p} \quad (67)$$

In using Eq. (67), we are faced with the problem of determining  $g(\mathbf{r}, \mathbf{p})$ , the number of states corresponding to each volume in phase space. Classical mechanics does not quantize variables, so there is no obvious way of doing

this. How can we make the connection between classical phase space and quantized states?<sup>14</sup> Having no reason to favor one region of phase space over another, we assume the following:

Equal volumes of phase space correspond to the same number of states, regardless of the specific values of  $\mathbf{r}$  and  $\mathbf{p}$  to which they apply.

We note that volumes in phase space have units of  $\text{kg m}^2/\text{s} = \text{J s}$  for each Cartesian coordinate. In order to have a dimensionless degeneracy,  $g$ , we divide volumes in phase space,  $d\mathbf{r} d\mathbf{p}$ , by a constant,  $h^3$ , for which  $h$  has units of  $\text{J s}$ . In other words,  $g(\mathbf{r}, \mathbf{p}) = 1h^{-3}$ , independent of position and momentum:

$$q = \frac{1}{h^3} \int \exp\left(-\frac{p^2}{2mkT}\right) d\mathbf{r} d\mathbf{p} \quad (68)$$

Because the three Cartesian directions are indistinguishable, Eq. (55) becomes

$$q = \frac{1}{h^3} \int \exp\left[-\frac{(p_x^2 + p_y^2 + p_z^2)}{2mkT}\right] dx dp_x dg dp_y dz dp_z = q_x^3 \quad (69)$$

where

$$q_x \equiv \frac{1}{h} \int_0^L dx \int_{-\infty}^{\infty} \exp\left(-\frac{p_x^2}{2mkT}\right) dp_x \quad (70)$$

From Eq. (54), the second integral is  $\sqrt{2\pi mkT}$ , giving<sup>15</sup>

$$q_x = \frac{L}{h} \sqrt{2\pi mkT} \quad (71)$$

or

$$q = \frac{V}{h^3} (2\pi mkT)^{3/2} \quad (72)$$

Using Eq. (26) to calculate the energy (realizing that the only term in  $\ln q$  that depends on  $T$  is  $\ln(T^{3/2})$ ):

$$U_m = U_0 + \frac{3}{2}RT \quad (73)$$

For  $C_{V,m}$ , this gives

$$C_{V,m} = \left(\frac{\partial U}{\partial T}\right)_V = \frac{3}{2}R \quad (74)$$

just the value that we obtained in [Chapter 2](#).

Quantum mechanics clearly denies the possibility of distinguishing between particles in translational motion.<sup>16</sup> For the ideal gas entropy, we must therefore use Eq. (34) for indistinguishable particles:

$$S_m = R \ln \left( \frac{V(2\pi mkT)^{3/2}}{N_A h^3} \right) + R + \frac{3}{2}R \quad (75)$$

Writing this in terms of the pressure for 1 mol gives

$$S_m = \frac{5}{2}R + R \ln \left( \frac{(2\pi M)^{3/2} R^{5/2} T^{5/2}}{N_A h^3 P} \right) \quad (76)$$

which is called the *Sackur–Tetrode equation*. Note that Eq. (76) gives the proper dependence for the entropy of an ideal gas on pressure at constant temperature [Eq. (28) of Chapter 4] and on temperature at constant pressure [Eq. (17) of Chapter 3]. In order to use Eq. (76) to calculate absolute entropies, a numerical value for  $h$  is required. This can be obtained by choosing  $h$  to fit the experimental absolute entropy for 1 mol of monatomic ideal gas (Problem 14). The result,  $h = 6.6 \times 10^{-34}$  J s, is called Planck's constant. It is the fundamental constant in the quantum description of matter. With the use of this constant, the Sackur–Tetrode equation gives good agreement with measured absolute entropies for monatomic gases at room temperature.

## 5.9 Energy of Polyatomic Ideal Gases

A polyatomic molecule contains more than one atom<sup>17</sup> and can undergo rotational and vibrational motions in addition to translational. Overall, rotational motion is unhindered and therefore consists only of a kinetic energy term of the form

$$\epsilon_{\text{rot}} = \frac{p_\theta^2}{2I} \quad (77)$$

where  $p_\theta$  is a rotational angular momentum and  $I$  is a moment of inertia. A vibrational motion in the harmonic approximation is of the form

$$\epsilon_{\text{vib}} = \frac{p_d^2}{2\mu} + \frac{k d^2}{2} \quad (78)$$

where  $d$  is the displacement of the vibrating oscillator from its equilibrium position and  $p_d$  is the momentum associated with vibration. ( $k$  and  $\mu$  are the force constant and mass, respectively, appropriate for the oscillator.) Of importance is

that in the harmonic approximation, the internal energy of a polyatomic molecule can be written as

$$\varepsilon_{\text{int}} = \sum_i C_i l_i^2 \quad (79)$$

where  $l_i$  is a coordinate *or* a momentum, and the sum goes over all of the coordinates of the rotations and the coordinates and momenta of the vibrations of the molecule.

Although vibrational and some rotational motions certainly require quantum mechanics for their accurate consideration, we will treat these motions in their classical limit. Using Eq. (23) and integrating over a  $2M$ -dimension phase space for a total of  $M$  rotations and vibrations, we get, using a normalization factor of  $h^{-3M}$ ,

$$\begin{aligned} q &= \frac{1}{h^{3M}} \int \exp\left(-\frac{1}{kT} \sum_i C_i l_i^2\right) d\mathbf{l}_j \\ &= \frac{1}{h^{3M}} \int d\mathbf{l}_{\text{rot}} \prod_i \int_{-\infty}^{\infty} \exp\left(-\frac{C_i}{kT} l_i^2\right) dl_i \end{aligned} \quad (80)$$

$\mathbf{l}_j$  includes a coordinate and momentum for each rotation and vibration, whereas  $l_i$  includes a coordinate and momentum for each vibration, but only a momentum for each rotation. Using Eq. (54) for each of the exponential integrals gives

$$q = \frac{V_\theta}{h^{3M}} \prod_i \sqrt{\frac{\pi kT}{C_i}} \quad (81)$$

where  $V_\theta$  is a factor of  $4\pi$  or  $8\pi^2$  obtained by integrating over the rotational coordinates. The final result does not depend on the value of this term. Inserting Eq. (81) into Eq. (26) gives

$$U_m = U_0 + RT^2 \left\{ \frac{\partial}{\partial T} \left[ \ln\left(\frac{V_\theta}{h^{3M}}\right) + \sum_i \ln\left(\frac{\pi k}{C_i}\right)^{1/2} + \sum_i \ln T^{1/2} \right]_V \right\} \quad (82)$$

Only the last term has a nonzero temperature derivative; therefore,

$$U_m = U_0 + \sum_i \frac{RT}{2} \quad (83)$$

This is the *classical equipartition theorem*. It states that each rotation (which only contributes one term to the sum) adds  $RT/2$  to the energy, whereas each vibration (which contributes two terms) adds  $RT$  to the energy. From Eq. (73), each of the

three translational motions of a molecule also contributes  $RT/2$  to the energy. In terms of heat capacities,

$$C_{V,m} = \left( \frac{\partial U_m}{\partial T} \right)_V = \frac{3}{2}R + \sum_{\text{rot}} \frac{1}{2}R + \sum_{\text{vib}} R \quad (84)$$

with an additional factor of  $R$  to calculate  $C_{P,m}$ . In calculating energies and heat capacities, one must enumerate the number of rotational and vibrational degrees of freedom of the molecule of interest. The key to this task is realizing that the same number of coordinates is required to localize the constituent atoms of the molecule, whether these atoms are bonded or not! If the molecule contains  $N$  atoms,  $3N$  coordinates (degrees of freedom) are required to localize these atoms in Cartesian space. When the atoms are bonded, we choose to use three coordinates to localize the center of mass of the ensemble of atoms. We call these the three translational degrees of freedom of the molecule. Likewise, the orientation of an axis of a linear molecule in space can be specified by two degrees of freedom, and if the molecule is nonlinear, an additional degree of freedom is required to specify the orientation of the molecule around this axis. The remaining  $3N - 6$  degrees of freedom ( $3N - 5$  in the case of linear molecules) must correspond to internal motions of the molecule, which we call vibrations.<sup>18</sup>

**Example 2.** Calculate the classical constant-pressure heat capacity of  $\text{CH}_4$ . Compare this with the literature value of  $35.46 \text{ J/mol K}$  at  $15^\circ\text{C}$ .

*Solution:*  $\text{CH}_4$  has 5 atoms, therefore  $N = 5$  and there are a total of  $3N = 15$  degrees of freedom. Three of these are translations, and because the molecule is nonlinear, three are rotations. There are, thus,  $15 - 6 = 9$  vibrational degrees of freedom. The internal energy is

$$U_m = \frac{3}{2}RT(\text{trans}) + \frac{3}{2}RT(\text{rot}) + 9RT(\text{vib}) = 12RT$$

and the heat capacities are  $C_{V,m} = (\partial U/\partial t)_V = 12R$  and

$$C_{P,m} = C_{V,m} + R = 13R = 108 \frac{\text{J}}{\text{mol K}}$$

This is over three times the literature value.

Real molecules follow quantum mechanics. However, the quantum results approach those of classical mechanics at high temperatures. “High temperature” is much lower for rotations and low-frequency vibrations than for high-frequency vibrations. For the latter, the molecule may dissociate before the classical approximation becomes appropriate. Comparing Eq. (84) for heat capacity with the data presented in Fig. 4 of Chapter 2, we note the following:

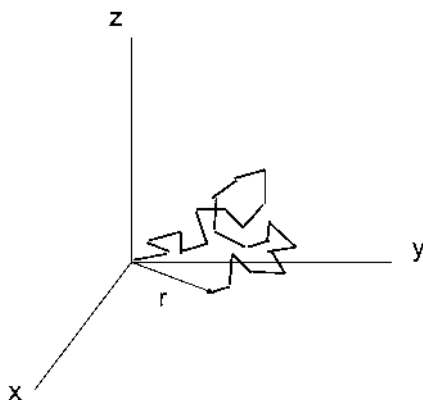


1. For monatomic gases, we have the complete translational contribution at all temperatures.
2. For diatomic gases, we have the translational and rotational (two degrees of freedom) contributions at room temperature, whereas vibrations only begin to contribute at higher temperature.
3. For polyatomic gases, some vibrational degrees of freedom start to contribute at room temperature, but the complete vibrational contribution is not even achieved at quite high temperature.

Observations that real energies and heat capacities are lower than those predicted by classical mechanics were one of a number of unsettling mysteries of physics that were cleared up by the development of quantum mechanics.

## 5.10 Configurations of a Polymer Chain

When we talk about a linear polymer chain, such as non-cross-linked polyethylene, we do not mean that all of the atoms in the chain are on a straight line or even that spatial structure repeats in some regular manner. Rather, the bond angles at the atoms in the chain are tetrahedral, and different conformations are obtained by rotation around the single bonds. A complete set of angles at all of the bonds gives a *configuration* of the chain, which can range from compact to extended. One particular configuration of a short polymer chain is shown in Fig. 2. We seek the components of the displacement vector,  $\mathbf{r}$ , between the ends of the chain. The magnitude of this vector,  $r$ , is called the *displacement length* and is a measure of the extension of the polymer chain.



**Figure 2** Configuration of a polymer chain.

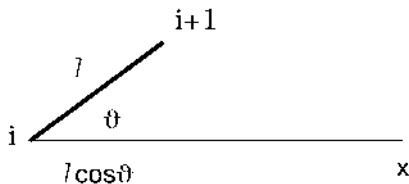
For ease of calculation, we make a number of simplifying assumptions. These are relaxed in advanced treatments of the subject. First, rather than requiring tetrahedral bonds at each vertex of the chain, we allow all bond angles and assume that these are randomly distributed. Second, we ignore any excluded volumes or interactions between the segments of the chain. In this sense, our calculation is similar to the Bernoulli model of the ideal gas, which neglects intermolecular interactions. Our approximation is called the *freely jointed chain* model.

We first consider a single bond in the chain. We want a measure of the projection of the length of this bond,  $l$ , in a direction, say  $x$ . The projection is shown in Fig. 3, with  $\theta$  varying from 0 to  $\pi$  radians. Because the average value of the projection is zero, we take its root-mean-square value as its measure. For a bond of length  $l$  at an angle of  $\theta$  with respect to the  $x$  axis, the projection is  $l \cos \theta$ . The root-mean-square (rms) projection is calculated by squaring this value, averaging it over all solid angles ( $d\Omega = 2\pi \sin \theta d\theta$ , and taking the square root of the result:

$$x_{\text{rms}} = \left( \frac{l^2}{4\pi} \int_0^\pi 2\pi \cos^2 \theta \sin \theta d\theta \right)^{1/2} = \frac{l}{\sqrt{3}} \quad (85)$$

Whereas  $x_{\text{rms}}$  is a measure of the magnitude of the projection of a single bond in the chain on the  $x$  axis, the projection can be in either the positive or negative direction. The actual length of the projection of the displacement vector on the  $x$  axis is proportional to the *difference* between the number of projections in the positive and negative  $x$  directions. Positive and negative projections are random and each has probability of  $\frac{1}{2}$ . For a chain having  $N$  segments, each of length  $l$ , the probability that the  $x$  projection of the extension vector has a value  $mx_{\text{rms}}$  is calculated in the following manner:

1. The probability of any particular ordering of positive and negative projections of the  $N$  bonds is  $(1/2)^N$ .
2. In order to have the difference between the number of positive projections and the number of negative projections equal  $m$ , with the



**Figure 3** Projection of a bond on the total extension vector.

total number of projections equal  $N$ , there must be  $(N + m)/2$  positive projections and  $(N - m)/2$  negative projections.

3. The number of ways of getting a difference of  $m$  out of a total of  $N$  projections is calculated by writing down all  $N!$  permutations of the segments and then dividing by the permutations of the positive and negative projections,

$$\left(\frac{N + m}{2}\right)! \left(\frac{N - m}{2}\right)!$$

because the order of these does not make a difference.

The probability is then

$$P(m, N) = \frac{N!}{[(N + m)/2]! [(N - m)/2]!} \left(\frac{1}{2}\right)^N \quad (86)$$

This probability distribution is called the (equal probability) *binomial distribution*, and is the same distribution that is obtained for tossing an unbiased coin. In books on statistics,<sup>19</sup> it is shown that for large values of  $N$ , the binomial distribution approaches the continuous *normal distribution*:

$$P(m, N) dm \approx \frac{1}{\sqrt{2\pi N}} \exp\left(-\frac{m^2}{2N}\right) dm \propto \Omega(m, N) \quad (87)$$

which is said to have the *Gaussian* form. We have indicated that the probability is proportional to the number of configurations that give rise to a chain with  $m$  and  $N$ . Substituting  $m = x/x_{\text{rms}} = x\sqrt{3}/l$  gives

$$P(x, N) dx = \sqrt{\frac{3}{2\pi Nl^2}} \exp\left(-\frac{3x^2}{2Nl^2}\right) dx \propto \Omega(x, N) dx \quad (88)$$

Similar distributions result for the projections along the  $y$  and  $z$  axes, and because these are independent,

$$\begin{aligned} P(x, y, z, N) dx dy dz &= \left(\frac{3}{2\pi Nl^2}\right)^{3/2} \exp\left(-\frac{3(x^2 + y^2 + z^2)}{2Nl^2}\right) dx dy dz \\ &\propto \Omega(x, y, z, N) dx dy dz \end{aligned} \quad (89)$$

To find the probability of a displacement length between  $r$  and  $r + dr$ , independent of direction, we integrate over the corresponding volume of space, much as we did in deriving the Maxwell–Boltzmann distribution:

$$P(r, N) dr = 4\pi \left(\frac{3}{2\pi Nl^2}\right)^{3/2} r^2 \exp\left(-\frac{3r^2}{2Nl^2}\right) dr \propto \Omega(r, N) dr \quad (90)$$

A single polymer chain is not a macroscopic system. However, for  $N$  of the order of 100 or more, the number of configurations is very large. We shall not hesitate to use our statistical thermodynamic equations in the mesoscopic range of single polymer chains.

## 5.11 Theory of Ideal Rubber Elasticity

Theories of different levels of complexity have been used to explain the elastic behavior of macromolecules (Fig. 12 of Chapter 1). The simplest theory, which nevertheless captures the essentials of the phenomenon, will be described here.<sup>20</sup> In this theory, a rubberlike polymer is pictured as a three-dimensional network of identical linear chains, each chain consisting of  $N$  bonds of length  $l$ . The chains are joined together at junctions, from which more than two chains extend. The junctions may be atoms chemically bonding more than two chains or tiny crystalline regions. In the language of polymer science, the chains are cross-linked and the cross-links prevent permanent deformation. The chains between the junctions are long and are assumed to adopt a random configuration by means of rotations around bonds. In the simplest theory, all rotational configurations are equally available; there are no excluded volume effects or intrachain interactions.<sup>21</sup> This is the model of the freely jointed chain, discussed in the previous section. All interactions between different chains are also neglected. The elasticity is completely intrachain and entropically produced.

Our object is to calculate a stress-strain relationship for comparison with Fig. 12 of Chapter 1. To do this, we calculate the force along the  $x$  direction necessary to distort a polymer sample, initially of dimensions  $X_0$ ,  $Y_0$ , and  $Z_0$  to dimensions  $X$ ,  $Y$ , and  $Z$ . We introduce distortion ratios,  $\alpha_x$ ,  $\alpha_y$ , and  $\alpha_z$  such that

$$X = \alpha_x X_0, \quad Y = \alpha_y Y_0, \quad Z = \alpha_z Z_0 \quad (91)$$

(In Chapter 4,  $\alpha_x$  was called  $\alpha$ , the elongation.) For stretching at constant volume, we have  $\alpha_x \alpha_y \alpha_z = 1$ , which, with  $\alpha_z = \alpha_y$ , gives

$$\alpha_y = \alpha_z = \frac{1}{\sqrt{\alpha_x}} \quad (92)$$

From Eq. (55) of Chapter 4 with  $X$  as the length of the polymer sample,

$$f = \left( \frac{\partial A}{\partial X} \right)_{T,V} = \left( \frac{\partial A}{\partial \alpha_x} \right)_{T,V} \frac{d\alpha_x}{dX} = \frac{1}{X_0} \left( \frac{\partial A}{\partial \alpha_x} \right)_{T,V} \quad (93)$$

Because  $A$ , the Helmholtz free energy, is an extensive property, it is given by

$$A = \kappa a \quad (94)$$

where  $a$  is the Helmholtz free energy of a single chain and  $\kappa$  is the number of chains in the sample:

$$f = \frac{\kappa}{X_0} \left( \frac{\partial a}{\partial \alpha_x} \right)_{T,V} = - \frac{\kappa T}{X_0} \left( \frac{\partial s}{\partial \alpha_x} \right)_{T,V} \quad (95)$$

where  $s$  is the entropy of a single chain.

The last equality follows for an ideal rubber, in which, by definition, the internal energy does not depend on the distortion.

We assume that the distortion ratios for the individual polymer chains are the same as those for the sample (affine approximation):

$$x = \alpha_x x_0, \quad y = \alpha_y y_0, \quad z = \alpha_z z_0 \quad (96)$$

(Lowercase symbols refer to the dimensions of individual polymer chains.) Also, in the absence of a force, the individual chains are isotropic:

$$x_0 = y_0 = z_0 \quad (97)$$

Using  $s = k \ln \Omega$  with Eq. (89),

$$s = - \frac{3k(\alpha_x^2 x_0^2 + \alpha_y^2 y_0^2 + \alpha_z^2 z_0^2)}{2Nl^2} \quad (98)$$

(Distortion-independent terms have been omitted.) Substituting Eqs. (92) and (97) into Eq. (98) gives

$$s = - \frac{3kx_0^2(\alpha_x^2 + 2/\alpha_x)}{2Nl^2} \quad (99)$$

Substituting into Eq. (95),

$$f = \frac{3kT\kappa x_0^2}{2Nl^2 X_0} (2\alpha_x - 2/\alpha_x^2). \quad (100)$$

Dividing by the area of the sample  $Y_0 Z_0$  and defining  $\kappa/X_0 Y_0 Z_0 \equiv \nu$ , the number of network chains per unit volume gives, for the stress,

$$f^* = \frac{3kT x_0^2 \nu}{Nl^2} (\alpha_x - 1/\alpha_x^2) \quad (101)$$

Because  $3x_0^2 = r_0^2$ , and  $r_0$ , the displacement length of a polymer chain is close to  $\sqrt{N}l$  (see Problem 19), Eq. (101) can be written

$$f^* \approx kT \nu (\alpha_x - 1/\alpha_x^2) \quad (102)$$

The simple theory presented accounts for the fall-off of the modulus of polymers with increased extension. The stiffening of the polymer at very large distortion results from interactions between the extended polymer chains, which is called *strain-induced crystallization*.

## Questions

1. Show that dividing Eq. (8) by  $N!$  produces no change in Eq. (10), and, thus, corrected Boltzmann statistics gives the same results as Boltzmann statistics.
2. Why is  $\int_{-\infty}^{\infty} x^{2n+1} \exp(-\beta x^2) dx = 0$ ?
3. Show that if the energy of a molecule can be written as the sum of terms for translational, rotational, and vibrational energies, the partition function for the molecule is the product of translational, rotational, and vibrational partition functions.
4. Why is the heat capacity of monatomic gases, such as He and Ne, practically independent of temperature, whereas molecular heat capacities increase with temperature?
5. Use Eq. (5) to discuss why coherent laser light has lower entropy than noncoherent monochromatic light and why a troop of marching soldiers has lower entropy than that of noonday pedestrians.
6. Give an example of a system containing distinguishable particles and one containing indistinguishable particles.
7. In Section 5.2, the probability of an observable spontaneous fluctuation away from the equilibrium state is given as  $1/\exp(N_A)$ . To get an idea of the magnitude of this quantity, compare  $\exp(N_A)$  with the number of atoms in the universe. You may make any assumptions that you like in calculating the latter quantity; you will see that they will make little difference in the comparison.
8. At a given temperature, which is larger: the molecular partition function,  $q$  or the system partition function,  $Q$ ?
9. What would happen if the two chambers maintained at different temperatures in Section 5.7 had small openings connecting them to the atmosphere and were therefore at the same pressure? Would the flow in both directions be the same and, if not, in which direction would there be a greater flow?
10. Compare Eq. (102) with the ideal gas equation. With reasonable values of  $v$  and  $\alpha$  could a stress of 1.0 atm be achieved?
11. Explain why the protein in egg white coagulates (becomes insoluble) when it is heated in boiling water.

## Problems

1. How many ways can you choose 3 students from a group of 10 students for the following.
  - (a) If the order in which you pick the students is important (corresponding to distinguishable molecules)
  - (b) If the order in which you pick the students is not important (corresponding to indistinguishable molecules)

2. How many configurations are there for the distribution of 10 distinguishable items in 3 states with occupation numbers  $N_0 = 5$ ,  $N_1 = 3$ , and  $N_2 = 2$ ?
3. How many configurations are there for four distinguishable molecules distributed over two states, with two molecules in each state? Write down these configurations.
4. Show that Eq. (4) for the reverse Joule process is unchanged if the particles are distinguishable.
- 5.\* Calculate the energy lost from a large container of monatomic gas if 1 mol of gas effuses out of a small hole in the wall of the container. Show that when some of the gas effuses out of the container, the remainder of the gas cools.
6. How accurate is Stirling's approximation, Eq. (28), for  $A = 3$  and  $A = 10$ , and  $A = 20$ ?
7. How many ways can you place 40 indistinguishable atoms in 1000 compartments?
8. Using Eq. (54), show that

$$\int_0^\infty \exp(-\beta x^2) dx = \frac{1}{2} \sqrt{\pi} \beta^{-1/2}$$

Note that  $0! = 1$ .

9. Verify Eq. (71).
10. Use the Maxwell–Boltzmann distribution to calculate  $c_{\text{rms}}$  and compare the result with that obtained from the Bernoulli model in [Chapter 1](#).
11. A single particle with energy states at 0, 1, and 2 has three states with energy below 2.5. How many states does a system consisting of three such particles have with energy below 2.5, for the following:
  - (a) The particles are distinguishable.
  - (b) The particles are indistinguishable.
12. Show that the Sackur–Tetrode equation [Eq. (76)] is in agreement with Eq. (17) of [Chapter 3](#) and Eq. (28) of [Chapter 4](#) for an ideal gas.
13. Use the Sackur–Tetrode equation [Eq. (76)] to calculate the standard molar entropy of neon considered an ideal gas at 298 K.
- 14.\* Use the Sackur–Tetrode equation [Eq. (76)] to evaluate  $h$  from the values of  $S_m^\circ$  of He and Ar at 298 K. (See Table 1 of Chapter 4.)
15. Calculate the classical  $C_{P,m}$  for Ar,  $\text{Cl}_2$ , and  $\text{CO}_2$  and compare these with the experimental values at 298 K for these gases, which are 20.93, 34.13, and 36.62 J/mol K, respectively.
- 16.\* Ice has a structure in which oxygen atoms are tetrahedrally arranged, and between every two oxygen atoms, there is a hydrogen atom. However, hydrogen atoms may be either covalently bonded or hydrogen-bonded to a given oxygen atom. Calculate the

number of possible configurations of the hydrogen atoms in 1 mol of ice in the following manner:

- (a) There are  $2N_A$  hydrogen atoms, each of which can be in two positions. How many configurations does this give?
- (b) However, not every configuration listed in part (a) is possible, because around each oxygen, there must be two covalently bonded and two hydrogen-bonded hydrogen atoms. For each of the  $N_A$  oxygen atoms, how many of the  $2^4 = 16$  arrangements of the hydrogen atoms give acceptable configurations?

Use your results from parts (a) and (b) to calculate the residual entropy of ice. The measured value is 3.4 J/mol K.

17. Derive the barometric distribution,  $P = P_0 \exp(-Mgh/RT)$ , discussed at the end of Chapter 1, starting from the Boltzmann distribution law [Eq. (16)].

18. Write down all of the possible results of four tosses of an unbiased coin. By counting, show that the probability for different numbers of “excess” heads is given by Eq. (86).

19. Find the most probable displacement length of a linear polymer chain, having 1000 bonds, each of length 0.14 nm.

20. Using Eqs. (55) and (90), show that the average displacement length of a freely jointed polymer chain with  $N$  bonds each of length  $d$  is within 10% of  $\sqrt{N}d$ .

21. The strain,  $\epsilon$ , on a rubber is  $\alpha_x - 1$ . Using this and Eq. (102), find expressions for  $a$  and  $b$  in Eq. (54) of Chapter 1. (Hint: You may want to expand a term in a power series in  $\epsilon$ .)

22.<sup>M</sup> Calculate the fraction of molecules that have speeds greater than twice the average molecular speed.

23. Find a formula for the most probable molecular speed,  $c_{mp}$ . Sketch the Maxwell–Boltzmann velocity distribution and show the relative positions of  $\langle c \rangle$ ,  $c_{mp}$ , and  $c_{rms}$  on your sketch.

24. Find a formula for the root-mean-cube velocity ( $c_{rmc} \equiv \sqrt[3]{\langle c^3 \rangle}$ ). Explain the trend of  $\langle c \rangle$ ,  $c_{rms}$ , and  $c_{rmc}$ .

## Notes

1. Quantum mechanics, the modern description of the microscopic world, also conserves energy.
2. In fact, quantum mechanics tells us that there must be uncertainty in the initial conditions and we can only discuss this process in a probabilistic manner.
3. We will assume that our camera could take the pictures without influencing the motion of the atoms. A real camera would observe light photons scattered off the atoms and these photons would influence the motion of the atoms.
4. See, for example, JE Mayer, MG Mayer. *Statistical Mechanics*. New York: Wiley, 1959, pp 53–63.



5. Quantum mechanics denies the possibility of exactly locating an atom or, more rigorously, of simultaneously determining both its position and momentum. In a combined position–momentum space (generally called phase space) quantum mechanics provides a “natural” size for the volume in which an atom can be located.
6. Logarithms to other bases, (e.g.,  $\log_{10}$ ) would also be an appropriate definition, and they are related to  $\ln$  by a constant multiplier.
7. This relationship, written as  $S = k \ln W$  is carved on Boltzmann’s gravestone in Vienna.
8. As a result of translational energy, energy is also often nearly continuous in quantum systems.
9. These types of particle are called fermions (electrons, protons, neutrons, and odd numbers of these) and bosons (photons, and even numbers of electrons, protons, and neutrons).
10. Quantum mechanical considerations indicate that it is necessary to keep the volume of the system constant in order for the energy levels of the system to be unchanged in the process.
11. This is the simplified form of Stirling’s approximation; there are more accurate (more complicated) forms.
12. We take  $U_0 = 0$  for the system.
13. We cannot avoid this by using a *gedanken* adiabatic wall between the two chambers, because molecules passing through the hole will transfer energy.
14. In quantum mechanics, *Heisenberg’s uncertainty principle* states that there is a limit to which we can know the *product* of the uncertainties in a coordinate and its corresponding momentum,  $\Delta x \Delta p_x$ . Thus, even in quantum mechanics, there is a minimum volume in phase space in which we can localize a particle.
15. We have assumed the gas to be confined in a cubic box with side  $L$ .
16. We can neither mark them (they are too small) nor watch them (the light bouncing off them will make this impossible).
17. Sometimes, diatomic molecules are distinguished from polyatomic molecules, which are defined to contain three or more atoms.
18. A molecule may also have internal rotations, which, if unhindered, should be counted as rotational motions, as no potential energy is involved in their motion.
19. For example, AJ Thomasian. *The Structure of Probability Theory with Applications*. New York: McGraw-Hill, 1969, p 253.
20. For a more exact development, see DA McQuarrie. *Statistical Thermodynamics*. New York: Harper and Row, 1973, pp 280–284.
21. More exactly, excluded volume effects and intrachain interactions are assumed not to change as the rubber is stretched.