

# 6 Boltzmann Statistics

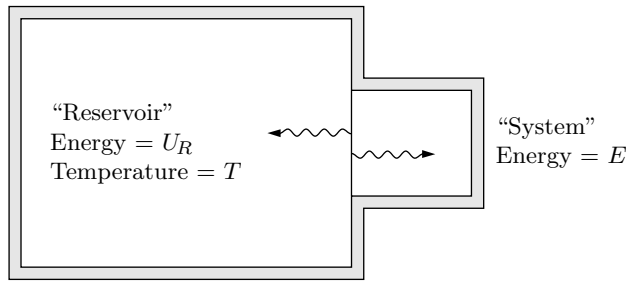
Most of this book so far has dealt with the second law of thermodynamics: its origin in the statistical behavior of large numbers of particles, and its applications in physics, chemistry, earth science, and engineering. However, the second law by itself usually doesn't tell us all we would like to know. In the last two chapters especially, we have often had to rely on experimental measurements (of enthalpies, entropies, and so on) before we could extract any predictions from the second law. This approach to thermodynamics can be extremely powerful, provided that the needed measurements can be made to the required precision.

Ideally, though, we would like to be able to calculate *all* thermodynamic quantities from first principles, starting from microscopic models of various systems of interest. In this book we have already worked with three important microscopic models: the two-state paramagnet, the Einstein solid, and the monatomic ideal gas. For each of these models we were able to write down an explicit combinatoric formula for the multiplicity,  $\Omega$ , and from there go on to compute the entropy, temperature, and other thermodynamic properties. In this chapter and the next two, we will study a number of more complicated models, representing a much greater variety of physical systems. For these more complicated models the direct combinatoric approach used in Chapters 2 and 3 would be too difficult mathematically. We therefore need to develop some new theoretical tools.

## 6.1 The Boltzmann Factor

In this section I will introduce the most powerful tool in all of statistical mechanics: an amazingly simple formula for the *probability* of finding a system in any particular microstate, when that system is in thermal equilibrium with a “reservoir” at a specified temperature (see Figure 6.1).

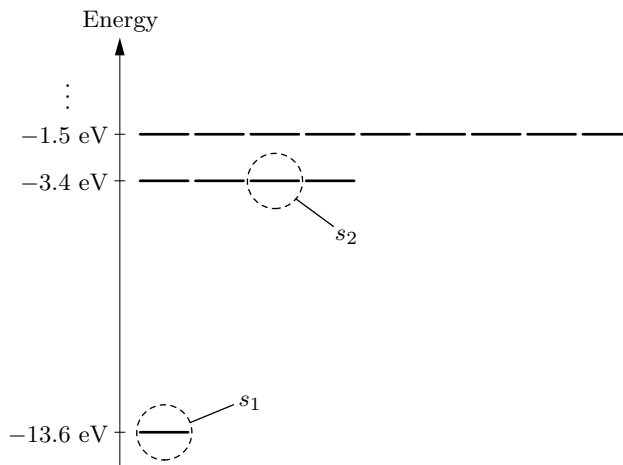
The system can be almost anything, but for definiteness, let's say it's a single atom. The microstates of the system then correspond to the various energy levels of



**Figure 6.1.** A “system” in thermal contact with a much larger “reservoir” at some well-defined temperature.

the atom, although for a given energy level there is often more than one independent state. For instance, a hydrogen atom has only one ground state (neglecting spin), with energy  $-13.6$  eV. But it has *four* independent states with energy  $-3.4$  eV, *nine* states with energy  $-1.5$  eV, and so on (see Figure 6.2). Each of these independent states counts as a separate microstate. When an energy level corresponds to more than one independent state, we say that level is **degenerate**. (For a more precise definition of degeneracy, and a more thorough discussion of the hydrogen atom, see Appendix A.)

If our atom were completely isolated from the rest of the universe, then its energy would be fixed, and all microstates with that energy would be equally probable. Now, however, we’re interested in the situation where the atom is not isolated, but instead is exchanging energy with lots of other atoms, which form a large “reservoir” with a fixed temperature. In this case the atom could conceivably be found in any of its microstates, but some will be more likely than others, depending on their energies. (Microstates with the *same* energy will still have the same probability.)



**Figure 6.2.** Energy level diagram for a hydrogen atom, showing the three lowest energy levels. There are four independent states with energy  $-3.4$  eV, and nine independent states with energy  $-1.5$  eV.

Since the probability of finding the atom in any particular microstate depends on how many other microstates there are, I'll simplify the problem, at first, by looking only at the *ratio* of probabilities for two particular microstates of interest (such as those circled in Figure 6.2). Let me call these states  $s_1$  and  $s_2$ , their energies  $E(s_1)$  and  $E(s_2)$ , and their probabilities  $\mathcal{P}(s_1)$  and  $\mathcal{P}(s_2)$ . How can I find a formula for the ratio of these probabilities? Let's go all the way back to the fundamental assumption of statistical mechanics: For an *isolated* system, all accessible microstates are equally probable. Our atom is not an isolated system, but the atom and the reservoir together do make an isolated system, and we are equally likely to find this combined system in any of its accessible microstates.

Now we don't care what the state of the reservoir is; we just want to know what state the atom is in. But if the atom is in state  $s_1$ , then the reservoir will have some very large number of accessible states, all equally probable. I'll call this number  $\Omega_R(s_1)$ : the multiplicity of the *reservoir* when the *atom* is in state  $s_1$ . Similarly, I'll use  $\Omega_R(s_2)$  to denote the multiplicity of the reservoir when the atom is in state  $s_2$ . These two multiplicities will generally be different, because when the atom is in a lower-energy state, more energy is left for the reservoir.

Suppose, for instance, that state  $s_1$  has the lower energy, so that  $\Omega_R(s_1) > \Omega_R(s_2)$ . As a specific example, let's say  $\Omega_R(s_1) = 100$  and  $\Omega_R(s_2) = 50$  (though in a true reservoir the multiplicities would be *much* larger). Now fundamentally, all microstates of the combined system are equally probable. Therefore since there are twice as many states of the combined system when the atom is in state  $s_1$  than when it is in state  $s_2$ , the former state must be twice as probable as the latter. More generally, the probability of finding the atom in any particular state is directly proportional to the number of microstates that are accessible to the reservoir. Thus the ratio of probabilities for any two states is

$$\frac{\mathcal{P}(s_2)}{\mathcal{P}(s_1)} = \frac{\Omega_R(s_2)}{\Omega_R(s_1)}. \quad (6.1)$$

Now we just have to get this expression into a more convenient form, using some math and a bit of thermodynamics.

First I'll rewrite each  $\Omega$  in terms of entropy, using the definition  $S = k \ln \Omega$ :

$$\frac{\mathcal{P}(s_2)}{\mathcal{P}(s_1)} = \frac{e^{S_R(s_2)/k}}{e^{S_R(s_1)/k}} = e^{[S_R(s_2) - S_R(s_1)]/k}. \quad (6.2)$$

The exponent now contains the *change* in the entropy of the reservoir, when the atom undergoes a transition from state 1 to state 2. This change will be tiny, since the atom is so small compared to the reservoir. We can therefore invoke the thermodynamic identity:

$$dS_R = \frac{1}{T} (dU_R + P dV_R - \mu dN_R). \quad (6.3)$$

The right-hand side involves the changes in the reservoir's energy, volume, and number of particles. But anything gained by the reservoir is lost by the atom, so

we can write each of these changes as minus the change in the same quantity for the atom.

I'm going to throw away the  $P dV$  and  $\mu dN$  terms, but for different reasons. The quantity  $P dV_R$  is usually nonzero, but much smaller than  $dU_R$  and therefore negligible. For instance, when an atom goes into an excited state, its effective volume might increase by a cubic ångström, so at atmospheric pressure, the term  $P dV$  is of order  $10^{-25}$  J. This is a million times less than the typical change in the atom's energy of a few electron-volts. Meanwhile,  $dN$  really is zero, at least when the small system is a single atom, and also in the other cases that we'll consider in this chapter. (In the following chapter I'll put the  $dN$  term back, in order to deal with other types of systems.)

So the difference of entropies in equation 6.2 can be rewritten

$$S_R(s_2) - S_R(s_1) = \frac{1}{T} [U_R(s_2) - U_R(s_1)] = -\frac{1}{T} [E(s_2) - E(s_1)], \quad (6.4)$$

where  $E$  is the energy of the atom. Plugging this expression back into equation 6.2, we obtain

$$\frac{\mathcal{P}(s_2)}{\mathcal{P}(s_1)} = e^{-[E(s_2) - E(s_1)]/kT} = \frac{e^{-E(s_2)/kT}}{e^{-E(s_1)/kT}}. \quad (6.5)$$

The ratio of probabilities is equal to a ratio of simple exponential factors, each of which is a function of the energy of the corresponding microstate and the temperature of the reservoir. Each of these exponential factors is called a **Boltzmann factor**:

$$\text{Boltzmann factor} = e^{-E(s)/kT}. \quad (6.6)$$

It would be nice if we could just say that the probability of each state is *equal* to the corresponding Boltzmann factor. Unfortunately, this isn't true. To arrive at the correct statement, let's manipulate equation 6.5 to get everything involving  $s_1$  on one side and everything involving  $s_2$  on the other side:

$$\frac{\mathcal{P}(s_2)}{e^{-E(s_2)/kT}} = \frac{\mathcal{P}(s_1)}{e^{-E(s_1)/kT}}. \quad (6.7)$$

Notice that the left side of this equation is independent of  $s_1$ ; therefore the right side must be as well. Similarly, since the right side is independent of  $s_2$ , so is the left side. But if both sides are independent of both  $s_1$  and  $s_2$ , then in fact both sides must be equal to a constant, the same for all states. The constant is called  $1/Z$ ; it is the constant of proportionality that converts a Boltzmann factor into a probability. In conclusion, for any state  $s$ ,

$$\mathcal{P}(s) = \frac{1}{Z} e^{-E(s)/kT}. \quad (6.8)$$

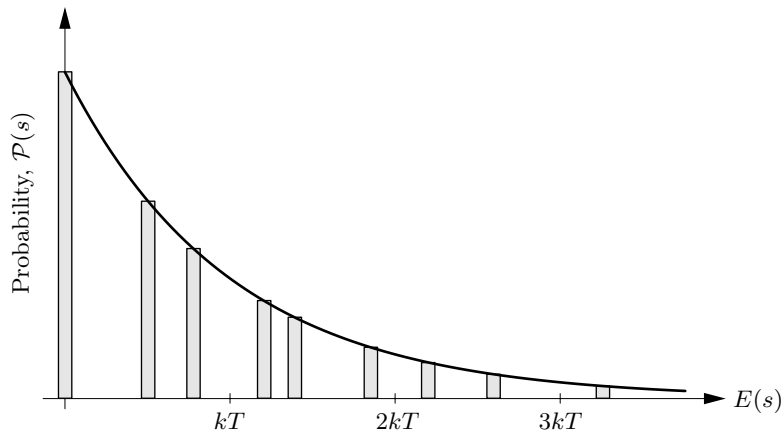
This is the most useful formula in all of statistical mechanics. Memorize it.\*

---

\*Equation 6.8 is sometimes called the **Boltzmann distribution** or the **canonical distribution**.

To interpret equation 6.8, let's suppose for a moment that the ground state energy of our atom is  $E_0 = 0$ , while the excited states all have positive energies. Then the probability of the ground state is  $1/Z$ , and all other states have smaller probabilities. States with energies much less than  $kT$  have probabilities only slightly less than  $1/Z$ , while states with energies much greater than  $kT$  have negligible probabilities, suppressed by the exponential Boltzmann factor. Figure 6.3 shows a bar graph of the probabilities for the states of a hypothetical system.

But what if the ground state energy is not zero? Physically, we should expect that shifting all the energies by a constant has no effect on the probabilities, and indeed, the probabilities *are* unchanged. It's true that all the Boltzmann factors get multiplied by an additional factor of  $e^{-E_0/kT}$ , but we'll see in a moment that  $Z$  gets multiplied by this same factor, so it cancels out in equation 6.8. Thus, the ground state still has the highest probability, and the remaining states have probabilities that are either a little less or a lot less, depending on how their energies, as measured from the ground state, compare to  $kT$ .



**Figure 6.3.** Bar graph of the relative probabilities of the states of a hypothetical system. The horizontal axis is energy. The smooth curve represents the Boltzmann distribution, equation 6.8, for one particular temperature. At lower temperatures it would fall off more suddenly, while at higher temperatures it would fall off more gradually.

**Problem 6.1.** Consider a system of two Einstein solids, where the first “solid” contains just a single oscillator, while the second solid contains 100 oscillators. The total number of energy units in the combined system is fixed at 500. Use a computer to make a table of the multiplicity of the combined system, for each possible value of the energy of the first solid from 0 units to 20. Make a graph of the total multiplicity vs. the energy of the first solid, and discuss, in some detail, whether the shape of the graph is what you would expect. Also plot the logarithm of the total multiplicity, and discuss the shape of this graph.

**Problem 6.2.** Prove that the probability of finding an atom in any particular energy level is  $\mathcal{P}(E) = (1/Z)e^{-F/kT}$ , where  $F = E - TS$  and the “entropy” of a level is  $k$  times the logarithm of the number of degenerate states for that level.

### The Partition Function

By now you're probably wondering how to actually *calculate*  $Z$ . The trick is to remember that the total probability of finding the atom in *some* state or other must be 1:

$$1 = \sum_s \mathcal{P}(s) = \sum_s \frac{1}{Z} e^{-E(s)/kT} = \frac{1}{Z} \sum_s e^{-E(s)/kT}. \quad (6.9)$$

Solving for  $Z$  therefore gives

$$Z = \sum_s e^{-E(s)/kT} = \text{sum of all Boltzmann factors}. \quad (6.10)$$

This sum isn't always easy to carry out, since there may be an infinite number of states  $s$  and you may not have a simple formula for their energies. But the terms in the sum get smaller and smaller as the energies  $E_s$  get larger, so often you can just compute the first several terms numerically, neglecting states with energies much greater than  $kT$ .

The quantity  $Z$  is called the **partition function**, and turns out to be far more useful than I would have suspected. It is a "constant" in that it does not depend on any particular state  $s$ , but it *does* depend on temperature. To interpret it further, suppose once again that the ground state has energy zero. Then the Boltzmann factor for the ground state is 1, and the rest of the Boltzmann factors are less than 1, by a little or a lot, in proportion to the probabilities of the associated states. Thus the partition function essentially *counts* how many states are accessible to the atom, weighting each one in proportion to its probability. At very low temperature,  $Z \approx 1$ , since all the excited states have very small Boltzmann factors. At high temperature,  $Z$  will be much larger. And if we shift all the energies by a constant  $E_0$ , the whole partition function just gets multiplied by an uninteresting factor of  $e^{-E_0/kT}$ , which cancels when we compute probabilities.

**Problem 6.3.** Consider a hypothetical atom that has just two states: a ground state with energy zero and an excited state with energy 2 eV. Draw a graph of the partition function for this system as a function of temperature, and evaluate the partition function numerically at  $T = 300$  K, 3000 K, 30,000 K, and 300,000 K.

**Problem 6.4.** Estimate the partition function for the hypothetical system represented in Figure 6.3. Then estimate the probability of this system being in its ground state.

**Problem 6.5.** Imagine a particle that can be in only three states, with energies  $-0.05$  eV, 0, and  $0.05$  eV. This particle is in equilibrium with a reservoir at 300 K.

- (a) Calculate the partition function for this particle.
- (b) Calculate the probability for this particle to be in each of the three states.
- (c) Because the zero point for measuring energies is arbitrary, we could just as well say that the energies of the three states are 0,  $+0.05$  eV, and  $+0.10$  eV, respectively. Repeat parts (a) and (b) using these numbers. Explain what changes and what doesn't.

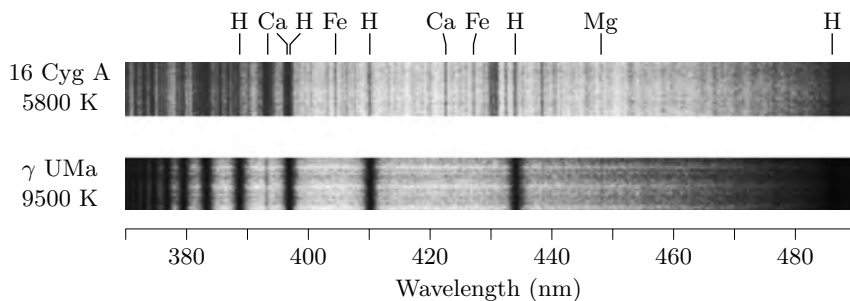
### Thermal Excitation of Atoms

As a simple application of Boltzmann factors, let's consider a hydrogen atom in the atmosphere of the sun, where the temperature is about 5800 K. (We'll see in the following chapter how this temperature can be measured from earth.) I'd like to compare the probability of finding the atom in one of its first excited states ( $s_2$ ) to the probability of finding it in the ground state ( $s_1$ ). The ratio of probabilities is the ratio of Boltzmann factors, so

$$\frac{\mathcal{P}(s_2)}{\mathcal{P}(s_1)} = \frac{e^{-E_2/kT}}{e^{-E_1/kT}} = e^{-(E_2-E_1)/kT}. \quad (6.11)$$

The difference in energy is 10.2 eV, while  $kT$  is  $(8.62 \times 10^{-5} \text{ eV/K})(5800 \text{ K}) = 0.50 \text{ eV}$ . So the ratio of probabilities is approximately  $e^{-20.4} = 1.4 \times 10^{-9}$ . For every billion atoms in the ground state, roughly 1.4 (on average) will be in any one of the first excited states. Since there are four such excited states, all with the same energy, the *total* number of atoms in these states will be four times as large, about 5.6 (for every billion in the ground state).

Atoms in the atmosphere of the sun can absorb sunlight on its way toward earth, but only at wavelengths that can induce transitions of the atoms into higher excited states. A hydrogen atom in its first excited state can absorb wavelengths in the Balmer series: 656 nm, 486 nm, 434 nm, and so on. These wavelengths are therefore missing, in part, from the sunlight we receive. If you put a narrow beam of sunlight through a good diffraction grating, you can see dark lines at the missing wavelengths (see Figure 6.4). There are also other prominent dark lines, created by other types of atoms in the solar atmosphere: iron, magnesium, sodium, calcium, and so on. The weird thing is, all these other wavelengths are absorbed by atoms



**Figure 6.4.** Photographs of the spectra of two stars. The upper spectrum is of a sunlike star (in the constellation Cygnus) with a surface temperature of about 5800 K; notice that the hydrogen absorption lines are clearly visible among a number of lines from other elements. The lower spectrum is of a hotter star (in Ursa Major, the Big Dipper), with a surface temperature of 9500 K. At this temperature a much larger fraction of the hydrogen atoms are in their first excited states, so the hydrogen lines are much more prominent than any others. Reproduced with permission from Helmut A. Abt et al., *An Atlas of Low-Dispersion Grating Stellar Spectra* (Kitt Peak National Observatory, Tucson, AZ, 1968).

(or ions) that start out either in their ground states or in very low-energy excited states (less than 3 eV above the ground state). The Balmer lines, by contrast, come only from the very rare hydrogen atoms that are excited more than 10 eV above the ground state. (A hydrogen atom in its ground state does not absorb any visible wavelengths.) Since the Balmer lines are quite prominent among the others, we can only conclude that hydrogen atoms are *much* more abundant in the sun's atmosphere than any of these other types.\*

**Problem 6.6.** Estimate the probability that a hydrogen atom at room temperature is in one of its first excited states (relative to the probability of being in the ground state). Don't forget to take degeneracy into account. Then repeat the calculation for a hydrogen atom in the atmosphere of the star  $\gamma$  UMa, whose surface temperature is approximately 9500 K.

**Problem 6.7.** Each of the hydrogen atom states shown in Figure 6.2 is actually twofold degenerate, because the electron can be in two independent spin states, both with essentially the same energy. Repeat the calculation given in the text for the relative probability of being in a first excited state, taking spin degeneracy into account. Show that the results are unaffected.

**Problem 6.8.** The energy required to ionize a hydrogen atom is 13.6 eV, so you might expect that the number of ionized hydrogen atoms in the sun's atmosphere would be even less than the number in the first excited state. Yet at the end of Chapter 5 I showed that the fraction of ionized hydrogen is much larger, nearly one atom in 10,000. Explain why this result is not a contradiction, and why it would be incorrect to try to calculate the fraction of ionized hydrogen using the methods of this section.

**Problem 6.9.** In the numerical example in the text, I calculated only the *ratio* of the probabilities of a hydrogen atom being in two different states. At such a low temperature the *absolute* probability of being in a first excited state is essentially the same as the relative probability compared to the ground state. Proving this rigorously, however, is a bit problematic, because a hydrogen atom has infinitely many states.

- (a) Estimate the partition function for a hydrogen atom at 5800 K, by adding the Boltzmann factors for all the states shown explicitly in Figure 6.2. (For simplicity you may wish to take the ground state energy to be zero, and shift the other energies accordingly.)
- (b) Show that if *all* bound states are included in the sum, then the partition function of a hydrogen atom is infinite, at any nonzero temperature. (See Appendix A for the full energy level structure of a hydrogen atom.)
- (c) When a hydrogen atom is in energy level  $n$ , the approximate radius of the electron wavefunction is  $a_0 n^2$ , where  $a_0$  is the Bohr radius, about  $5 \times 10^{-11}$  m. Going back to equation 6.3, argue that the  $P dV$  term is *not* negligible for the very high- $n$  states, and therefore that the result of part (a), not that of part (b), gives the physically relevant partition function for this problem. Discuss.

---

\*The recipe of the stars was first worked out by Cecilia Payne in 1924. The story is beautifully told by Philip and Phylis Morrison in *The Ring of Truth* (Random House, New York, 1987).



**Problem 6.10.** A water molecule can vibrate in various ways, but the easiest type of vibration to excite is the “flexing” mode in which the hydrogen atoms move toward and away from each other but the HO bonds do not stretch. The oscillations of this mode are approximately harmonic, with a frequency of  $4.8 \times 10^{13}$  Hz. As for any quantum harmonic oscillator, the energy levels are  $\frac{1}{2}hf$ ,  $\frac{3}{2}hf$ ,  $\frac{5}{2}hf$ , and so on. None of these levels are degenerate.

- (a) Calculate the probability of a water molecule being in its flexing ground state and in each of the first two excited states, assuming that it is in equilibrium with a reservoir (say the atmosphere) at 300 K. (Hint: Calculate  $Z$  by adding up the first few Boltzmann factors, until the rest are negligible.)
- (b) Repeat the calculation for a water molecule in equilibrium with a reservoir at 700 K (perhaps in a steam turbine).

**Problem 6.11.** A lithium nucleus has four independent spin orientations, conventionally labeled by the quantum number  $m = -3/2, -1/2, 1/2, 3/2$ . In a magnetic field  $B$ , the energies of these four states are  $E = -m\mu B$ , where the constant  $\mu$  is  $1.03 \times 10^{-7}$  eV/T. In the Purcell-Pound experiment described in Section 3.3, the maximum magnetic field strength was 0.63 T and the temperature was 300 K. Calculate the probability of a lithium nucleus being in each of its four spin states under these conditions. Then show that, if the field is suddenly reversed, the probabilities of the four states obey the Boltzmann distribution for  $T = -300$  K.

**Problem 6.12.** Cold interstellar molecular clouds often contain the molecule cyanogen (CN), whose first rotational excited states have an energy of  $4.7 \times 10^{-4}$  eV (above the ground state). There are actually three such excited states, all with the same energy. In 1941, studies of the absorption spectrum of starlight that passes through these molecular clouds showed that for every ten CN molecules that are in the ground state, approximately three others are in the three first excited states (that is, an average of one in each of these states). To account for this data, astronomers suggested that the molecules might be in thermal equilibrium with some “reservoir” with a well-defined temperature. What is that temperature?\*

**Problem 6.13.** At *very* high temperatures (as in the very early universe), the proton and the neutron can be thought of as two different states of the same particle, called the “nucleon.” (The reactions that convert a proton to a neutron or vice versa require the absorption of an electron or a positron or a neutrino, but all of these particles tend to be very abundant at sufficiently high temperatures.) Since the neutron’s mass is higher than the proton’s by  $2.3 \times 10^{-30}$  kg, its energy is higher by this amount times  $c^2$ . Suppose, then, that at some very early time, the nucleons were in thermal equilibrium with the rest of the universe at  $10^{11}$  K. What fraction of the nucleons at that time were protons, and what fraction were neutrons?

**Problem 6.14.** Use Boltzmann factors to derive the exponential formula for the density of an isothermal atmosphere, already derived in Problems 1.16 and 3.37. (Hint: Let the system be a single air molecule, let  $s_1$  be a state with the molecule at sea level, and let  $s_2$  be a state with the molecule at height  $z$ .)

---

\*For a review of these measurements and calculations, see Patrick Thaddeus, *Annual Reviews of Astronomy and Astrophysics* **10**, 305–334 (1972).

## 6.2 Average Values

In the previous section we saw how to calculate the probability that a system is in any particular one of its microstates  $s$ , given that it is in equilibrium with a reservoir at temperature  $T$ :

$$\mathcal{P}(s) = \frac{1}{Z} e^{-\beta E(s)}, \quad (6.12)$$

where  $\beta$  is an abbreviation for  $1/kT$ . The exponential factor is called the **Boltzmann factor**, while  $Z$  is the **partition function**,

$$Z = \sum_s e^{-\beta E(s)}, \quad (6.13)$$

that is, the sum of the Boltzmann factors for all possible states.

Suppose, though, that we're not interested in knowing all the probabilities of all the various states our system could be in—suppose we just want to know the *average* value of some property of the system, such as its energy. Is there an easy way to compute this average, and if so, how?

Let me give a simple example. Suppose my system is an atom that has just three possible states: The ground state with energy 0 eV, a state with energy 4 eV, and a state with energy 7 eV. Actually, though, I have five such atoms, and at the moment, two of them are in the ground state, two are in the 4-eV state, and one is in the 7-eV state (see Figure 6.5). What is the average energy of all my atoms? Just add 'em up and divide by 5:

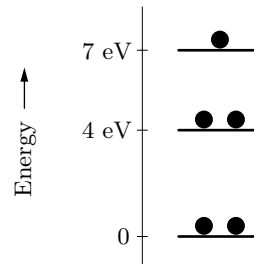
$$\overline{E} = \frac{(0 \text{ eV}) \cdot 2 + (4 \text{ eV}) \cdot 2 + (7 \text{ eV}) \cdot 1}{5} = 3 \text{ eV}. \quad (6.14)$$

But there's another way to think about this computation. Instead of computing the numerator first and then dividing by 5, we can group the  $1/5$  with the factors of 2, 2, and 1, which represent the numbers of atoms in each state:

$$\overline{E} = (0 \text{ eV}) \cdot \frac{2}{5} + (4 \text{ eV}) \cdot \frac{2}{5} + (7 \text{ eV}) \cdot \frac{1}{5} = 3 \text{ eV}. \quad (6.15)$$

In this expression, the energy of each state is multiplied by the *probability* of that state occurring (in any particular atom chosen from among my sample of 5); those probabilities are just  $2/5$ ,  $2/5$ , and  $1/5$ , respectively.

**Figure 6.5.** Five hypothetical atoms distributed among three different states.



It's not hard to generalize this example into a formula. If I have a large sample of  $N$  atoms, and  $N(s)$  is the number of atoms in any particular state  $s$ , then the average value of the energy is

$$\overline{E} = \frac{\sum_s E(s)N(s)}{N} = \sum_s E(s) \frac{N(s)}{N} = \sum_s E(s)\mathcal{P}(s), \quad (6.16)$$

where  $\mathcal{P}(s)$  is the probability of finding an atom in state  $s$ . So the average energy is just the sum of all the energies, weighted by their probabilities.

In the statistical mechanical systems that we're considering, each probability is given by equation 6.12, so

$$\overline{E} = \frac{1}{Z} \sum_s E(s) e^{-\beta E(s)}. \quad (6.17)$$

Notice that the sum is similar to the partition function (6.13), but with an extra factor of  $E(s)$  in each term.\*

The average value of any other variable of interest can be computed in exactly the same way. If the variable is called  $X$ , and has the value  $X(s)$  in state  $s$ , then

$$\overline{X} = \sum_s X(s)\mathcal{P}(s) = \frac{1}{Z} \sum_s X(s) e^{-\beta E(s)}. \quad (6.18)$$

One nice feature of average values is that they are additive; for example, the average total energy of two objects is the sum of their individual average energies. This means that if you have a collection of many identical, independent particles, you can compute their total (average) energy from the average energy of just one, simply by multiplying by how many there are:

$$U = N\overline{E}. \quad (6.19)$$

(Now you see why I've been using the symbol  $E$  for the energy of the atom; I've reserved  $U$  for the total energy of the much larger system that contains it.) So when I divided things into an "atom" and a "reservoir" in the previous section, it was partly just a trick. Even if you want to know the total energy of the *whole* system, you can often find it by concentrating first on one particle in the system, treating the rest as the reservoir. Once you know the average value of the quantity of interest for your particle, just multiply by  $N$  to get the total.

Technically, the  $U$  in equation 6.19 is merely the *average* energy of the entire system. If even this large system is in thermal contact with other objects, then the instantaneous value of  $U$  will fluctuate away from the average. However, if  $N$  is large, these fluctuations will almost always be negligible. Problem 6.17 shows how to calculate the size of typical fluctuations.

---

\*In this chapter, the set of systems that we average *over* will be a hypothetical set whose members are assigned to states according to the Boltzmann probability distribution. This hypothetical set of systems is often called a **canonical ensemble**. In Chapters 2 and 3 we instead worked with isolated systems, where all allowed states had the same probability; a set of hypothetical systems with that (trivial) probability distribution is called a **microcanonical ensemble**.

**Problem 6.15.** Suppose you have 10 atoms of weberium: 4 with energy 0 eV, 3 with energy 1 eV, 2 with energy 4 eV, and 1 with energy 6 eV.

- Compute the average energy of all your atoms, by adding up all their energies and dividing by 10.
- Compute the probability that one of your atoms chosen at random would have energy  $E$ , for each of the four values of  $E$  that occur.
- Compute the average energy again, using the formula  $\bar{E} = \sum_s E(s)\mathcal{P}(s)$ .

**Problem 6.16.** Prove that, for any system in equilibrium with a reservoir at temperature  $T$ , the average value of the energy is

$$\bar{E} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial}{\partial \beta} \ln Z,$$

where  $\beta = 1/kT$ . These formulas can be extremely useful when you have an explicit formula for the partition function.

**Problem 6.17.** The most common measure of the *fluctuations* of a set of numbers away from the average is the **standard deviation**, defined as follows.

- For each atom in the five-atom toy model of Figure 6.5, compute the *deviation* of the energy from the average energy, that is,  $E_i - \bar{E}$ , for  $i = 1$  to 5. Call these deviations  $\Delta E_i$ .
- Compute the average of the *squares* of the five deviations, that is,  $\overline{(\Delta E_i)^2}$ . Then compute the square root of this quantity, which is the root-mean-square (rms) deviation, or standard deviation. Call this number  $\sigma_E$ . Does  $\sigma_E$  give a reasonable measure of how far the individual values tend to stray from the average?
- Prove in general that

$$\sigma_E^2 = \overline{E^2} - (\bar{E})^2,$$

that is, the standard deviation squared is the average of the squares minus the square of the average. This formula usually gives the easier way of computing a standard deviation.

- Check the preceding formula for the five-atom toy model of Figure 6.5.

**Problem 6.18.** Prove that, for any system in equilibrium with a reservoir at temperature  $T$ , the average value of  $E^2$  is

$$\overline{E^2} = \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2}.$$

Then use this result and the results of the previous two problems to derive a formula for  $\sigma_E$  in terms of the heat capacity,  $C = \partial \bar{E} / \partial T$ . You should find

$$\sigma_E = kT \sqrt{C/k}.$$

**Problem 6.19.** Apply the result of Problem 6.18 to obtain a formula for the standard deviation of the energy of a system of  $N$  identical harmonic oscillators (such as in an Einstein solid), in the high-temperature limit. Divide by the average energy to obtain a measure of the *fractional* fluctuation in energy. Evaluate this fraction numerically for  $N = 1$ ,  $10^4$ , and  $10^{20}$ . Discuss the results briefly.

### Paramagnetism

As a first application of these tools, I'd like to rederive some of our earlier results (see Section 3.3) for the ideal two-state paramagnet.

Recall that each elementary dipole in an ideal two-state paramagnet has just two possible states: an “up” state with energy  $-\mu B$ , and a “down” state with energy  $+\mu B$ . (Here  $B$  is the strength of the externally applied magnetic field, while the component of the dipole's magnetic moment in the direction of the field is  $\pm\mu$ .) The partition function for a single dipole is therefore

$$Z = \sum_s e^{-\beta E(s)} = e^{+\beta\mu B} + e^{-\beta\mu B} = 2 \cosh(\beta\mu B). \quad (6.20)$$

The probability of finding the dipole in the “up” state is

$$\mathcal{P}_\uparrow = \frac{e^{+\beta\mu B}}{Z} = \frac{e^{+\beta\mu B}}{2 \cosh(\beta\mu B)}, \quad (6.21)$$

while the probability of finding it in the “down” state is

$$\mathcal{P}_\downarrow = \frac{e^{-\beta\mu B}}{Z} = \frac{e^{-\beta\mu B}}{2 \cosh(\beta\mu B)}. \quad (6.22)$$

You can easily check that these two probabilities add up to 1.

The average energy of our dipole is

$$\begin{aligned} \bar{E} &= \sum_s E(s) \mathcal{P}(s) = (-\mu B) \mathcal{P}_\uparrow + (+\mu B) \mathcal{P}_\downarrow = -\mu B (\mathcal{P}_\uparrow - \mathcal{P}_\downarrow) \\ &= -\mu B \frac{e^{\beta\mu B} - e^{-\beta\mu B}}{2 \cosh(\beta\mu B)} = -\mu B \tanh(\beta\mu B). \end{aligned} \quad (6.23)$$

If we have a collection of  $N$  such dipoles, the total energy is

$$U = -N\mu B \tanh(\beta\mu B), \quad (6.24)$$

in agreement with equation 3.31. In Section 3.3, however, we had to work much harder to derive this result: We started with the exact combinatoric formula for the multiplicity, then applied Stirling's approximation to simplify the entropy, then took a derivative and did lots of algebra to finally get  $U$  as a function of  $T$ . Here all we needed was Boltzmann factors.

According to the result of Problem 6.16, we can also compute the average energy by differentiating  $Z$  with respect to  $\beta$ , then multiplying by  $-1/Z$ :

$$\bar{E} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta}. \quad (6.25)$$

Let's check this formula for the two-state paramagnet:

$$\bar{E} = -\frac{1}{Z} \frac{\partial}{\partial \beta} 2 \cosh(\beta\mu B) = -\frac{1}{Z} (2\mu B) \sinh(\beta\mu B) = -\mu B \tanh(\beta\mu B). \quad (6.26)$$

Yep, it works.

Finally, we can compute the average value of a dipole's magnetic moment along the direction of  $B$ :

$$\overline{\mu_z} = \sum_s \mu_z(s) \mathcal{P}(s) = (+\mu)\mathcal{P}_\uparrow + (-\mu)\mathcal{P}_\downarrow = \mu \tanh(\beta\mu B). \quad (6.27)$$

Thus the total magnetization of the sample is

$$M = N\overline{\mu_z} = N\mu \tanh(\beta\mu B), \quad (6.28)$$

in agreement with equation 3.32.

**Problem 6.20.** This problem concerns a collection of  $N$  identical harmonic oscillators (perhaps an Einstein solid or the internal vibrations of gas molecules) at temperature  $T$ . As in Section 2.2, the allowed energies of each oscillator are 0,  $hf$ ,  $2hf$ , and so on.

(a) Prove by long division that

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots.$$

For what values of  $x$  does this series have a finite sum?

- (b) Evaluate the partition function for a single harmonic oscillator. Use the result of part (a) to simplify your answer as much as possible.
- (c) Use formula 6.25 to find an expression for the average energy of a single oscillator at temperature  $T$ . Simplify your answer as much as possible.
- (d) What is the total energy of the system of  $N$  oscillators at temperature  $T$ ? Your result should agree with what you found in Problem 3.25.
- (e) If you haven't already done so in Problem 3.25, compute the heat capacity of this system and check that it has the expected limits as  $T \rightarrow 0$  and  $T \rightarrow \infty$ .

**Problem 6.21.** In the real world, most oscillators are not perfectly harmonic. For a quantum oscillator, this means that the spacing between energy levels is not exactly uniform. The vibrational levels of an  $\text{H}_2$  molecule, for example, are more accurately described by the approximate formula

$$E_n \approx \epsilon(1.03n - 0.03n^2), \quad n = 0, 1, 2, \dots,$$

where  $\epsilon$  is the spacing between the two lowest levels. Thus, the levels get closer together with increasing energy. (This formula is reasonably accurate only up to about  $n = 15$ ; for slightly higher  $n$  it would say that  $E_n$  decreases with increasing  $n$ . In fact, the molecule dissociates and there are no more discrete levels beyond  $n \approx 15$ .) Use a computer to calculate the partition function, average energy, and heat capacity of a system with this set of energy levels. Include all levels through  $n = 15$ , but check to see how the results change when you include fewer levels. Plot the heat capacity as a function of  $kT/\epsilon$ . Compare to the case of a perfectly harmonic oscillator with evenly spaced levels, and also to the vibrational portion of the graph in Figure 1.13.

**Problem 6.22.** In most paramagnetic materials, the individual magnetic particles have more than two independent states (orientations). The number of independent states depends on the particle's angular momentum “quantum number”  $j$ , which must be a multiple of  $1/2$ . For  $j = 1/2$  there are just two independent states, as discussed in the text above and in Section 3.3. More generally, the allowed values of the  $z$  component of a particle's magnetic moment are

$$\mu_z = -j\delta_\mu, (-j+1)\delta_\mu, \dots, (j-1)\delta_\mu, j\delta_\mu,$$

where  $\delta_\mu$  is a constant, equal to the difference in  $\mu_z$  between one state and the next. (When the particle's angular momentum comes entirely from electron spins,  $\delta_\mu$  equals twice the Bohr magneton. When orbital angular momentum also contributes,  $\delta_\mu$  is somewhat different but comparable in magnitude. For an atomic nucleus,  $\delta_\mu$  is roughly a thousand times smaller.) Thus the number of states is  $2j+1$ . In the presence of a magnetic field  $B$  pointing in the  $z$  direction, the particle's magnetic energy (neglecting interactions between dipoles) is  $-\mu_z B$ .

- (a) Prove the following identity for the sum of a finite geometric series:

$$1 + x + x^2 + \dots + x^n = \frac{1 - x^{n+1}}{1 - x}.$$

(Hint: Either prove this formula by induction on  $n$ , or write the series as a difference between two infinite series and use the result of Problem 6.20(a).)

- (b) Show that the partition function of a single magnetic particle is

$$Z = \frac{\sinh[b(j + \frac{1}{2})]}{\sinh \frac{b}{2}},$$

where  $b = \beta\delta_\mu B$ .

- (c) Show that the total magnetization of a system of  $N$  such particles is

$$M = N\delta_\mu \left[ (j + \frac{1}{2}) \coth[b(j + \frac{1}{2})] - \frac{1}{2} \coth \frac{b}{2} \right],$$

where  $\coth x$  is the hyperbolic cotangent, equal to  $\cosh x / \sinh x$ . Plot the quantity  $M/N\delta_\mu$  vs.  $b$ , for a few different values of  $j$ .

- (d) Show that the magnetization has the expected behavior as  $T \rightarrow 0$ .  
 (e) Show that the magnetization is proportional to  $1/T$  (Curie's law) in the limit  $T \rightarrow \infty$ . (Hint: First show that  $\coth x \approx \frac{1}{x} + \frac{x}{3}$  when  $x \ll 1$ .)  
 (f) Show that for  $j = 1/2$ , the result of part (c) reduces to the formula derived in the text for a two-state paramagnet.

## Rotation of Diatomic Molecules

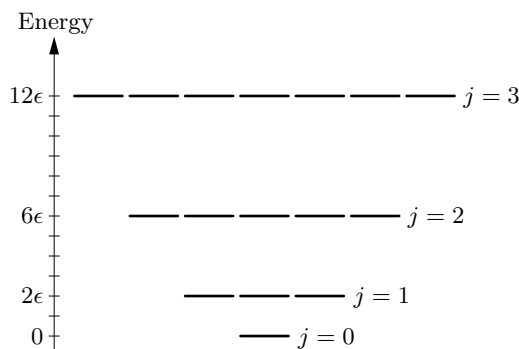
Now let's consider a more intricate application of Boltzmann factors and average values: the rotational motion of a diatomic molecule (assumed to be isolated, as in a low-density gas).

Rotational energies are quantized. (For details, see Appendix A.) For a diatomic molecule like CO or HCl, the allowed rotational energies are

$$E(j) = j(j+1)\epsilon, \quad (6.29)$$

where  $j$  can be 0, 1, 2, etc., and  $\epsilon$  is a constant that is inversely proportional to the molecule's moment of inertia. The number of degenerate states for level  $j$  is  $2j+1$ ,

**Figure 6.6.** Energy level diagram for the rotational states of a diatomic molecule.



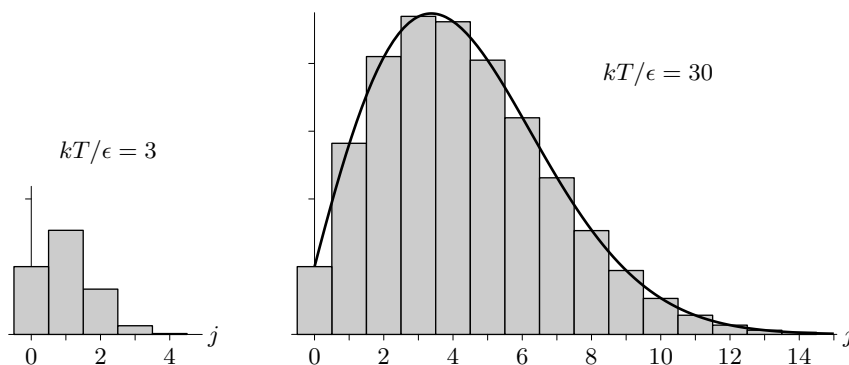
as shown in Figure 6.6. (I'm assuming, for now, that the two atoms making up the molecule are of different types. For molecules made of identical atoms, like  $\text{H}_2$  or  $\text{N}_2$ , there is a subtlety that I'll deal with later.)

Given this energy level structure, we can write the partition function as a sum over  $j$ :

$$Z_{\text{rot}} = \sum_{j=0}^{\infty} (2j+1)e^{-E(j)/kT} = \sum_{j=0}^{\infty} (2j+1)e^{-j(j+1)\epsilon/kT}. \quad (6.30)$$

Figure 6.7 shows a pictorial representation of this sum as the area under a bar graph. Unfortunately, there is no way to evaluate the sum exactly in closed form. But it's not hard to evaluate the sum numerically, for any particular temperature. Even better, in most cases of interest we can approximate the sum as an integral that yields a very simple result.

Let's look at some numbers. The constant  $\epsilon$ , which sets the energy scale for rotational excitations, is never more than a small fraction of an electron-volt. For a CO molecule, for instance,  $\epsilon = 0.00024$  eV, so that  $\epsilon/k = 2.8$  K. Ordinarily we are interested only in temperatures much higher than  $\epsilon/k$ , so the quantity  $kT/\epsilon$  will be much greater than 1. In this case the number of terms that contribute significantly



**Figure 6.7.** Bar-graph representations of the partition sum 6.30, for two different temperatures. At high temperatures the sum can be approximated as the area under a smooth curve.



to the partition function will be quite large, so we can, to a good approximation, replace the bar graph in Figure 6.7 with the smooth curve. The partition function is then approximately the area under this curve:

$$Z_{\text{rot}} \approx \int_0^\infty (2j+1)e^{-j(j+1)\epsilon/kT} dj = \frac{kT}{\epsilon} \quad (\text{when } kT \gg \epsilon). \quad (6.31)$$

(To evaluate the integral, make the substitution  $x = j(j+1)\epsilon/kT$ .) This result should be accurate in the high-temperature limit where  $Z_{\text{rot}} \gg 1$ . As expected, the partition function increases with increasing temperature. For CO at room temperature,  $Z_{\text{rot}}$  is slightly greater than 100 (see Problem 6.23).

Still working in the high-temperature approximation, we can calculate the average rotational energy of a molecule using the magical formula 6.25:

$$\overline{E}_{\text{rot}} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -(\beta\epsilon) \frac{\partial}{\partial \beta} \frac{1}{\beta\epsilon} = \frac{1}{\beta} = kT \quad (\text{when } kT \gg \epsilon). \quad (6.32)$$

This is just the prediction of the equipartition theorem, since a diatomic molecule has two rotational degrees of freedom. Differentiating  $\overline{E}$  with respect to  $T$  gives the contribution of this energy to the heat capacity, simply  $k$  (for each molecule), again in agreement with the equipartition theorem. At low temperature, however, the third law tells us that the heat capacity must go to *zero*; and indeed it does, as you can confirm from the exact expression 6.30 (see Problem 6.26).

So much for diatomic molecules made of distinguishable atoms. Now, what about the case of identical atoms, such as the important molecules  $\text{N}_2$  and  $\text{O}_2$ ? The subtlety here is that turning the molecule by  $180^\circ$  does not change its spatial configuration, so the molecule actually has only *half* as many states as it otherwise would. In the high-temperature limit, when  $Z \gg 1$ , we can account for this symmetry by inserting a factor of  $1/2$  into the partition function:

$$Z_{\text{rot}} \approx \frac{kT}{2\epsilon} \quad (\text{identical atoms, } kT \gg \epsilon). \quad (6.33)$$

The factor of  $1/2$  cancels out of the average energy (equation 6.32), so it has no effect on the heat capacity. At lower temperatures, however, things become more complicated: One must figure out exactly which terms should be omitted from the partition function (equation 6.30). At ordinary pressures, all diatomic gases except hydrogen will liquefy long before such low temperatures are reached. The behavior of hydrogen at low temperature is the subject of Problem 6.30.

**Problem 6.23.** For a CO molecule, the constant  $\epsilon$  is approximately 0.00024 eV. (This number is measured using microwave spectroscopy, that is, by measuring the microwave frequencies needed to excite the molecules into higher rotational states.) Calculate the rotational partition function for a CO molecule at room temperature (300 K), first using the exact formula 6.30 and then using the approximate formula 6.31.

**Problem 6.24.** For an  $\text{O}_2$  molecule, the constant  $\epsilon$  is approximately 0.00018 eV. Estimate the rotational partition function for an  $\text{O}_2$  molecule at room temperature.

**Problem 6.25.** The analysis of this section applies also to linear polyatomic molecules, for which no rotation about the axis of symmetry is possible. An example is  $\text{CO}_2$ , with  $\epsilon = 0.000049$  eV. Estimate the rotational partition function for a  $\text{CO}_2$  molecule at room temperature. (Note that the arrangement of the atoms is OCO, and the two oxygen atoms are identical.)

**Problem 6.26.** In the *low*-temperature limit ( $kT \ll \epsilon$ ), each term in the rotational partition function (equation 6.30) is much smaller than the one before. Since the first term is independent of  $T$ , cut off the sum after the second term and compute the average energy and the heat capacity in this approximation. Keep only the largest  $T$ -dependent term at each stage of the calculation. Is your result consistent with the third law of thermodynamics? Sketch the behavior of the heat capacity at all temperatures, interpolating between the high-temperature and low-temperature expressions.

**Problem 6.27.** Use a computer to sum the exact rotational partition function (equation 6.30) numerically, and plot the result as a function of  $kT/\epsilon$ . Keep enough terms in the sum to be confident that the series has converged. Show that the approximation in equation 6.31 is a bit low, and estimate by how much. Explain the discrepancy by referring to Figure 6.7.

**Problem 6.28.** Use a computer to sum the rotational partition function (equation 6.30) algebraically, keeping terms through  $j = 6$ . Then calculate the average energy and the heat capacity. Plot the heat capacity for values of  $kT/\epsilon$  ranging from 0 to 3. Have you kept enough terms in  $Z$  to give accurate results within this temperature range?

**Problem 6.29.** Although an ordinary  $\text{H}_2$  molecule consists of two identical atoms, this is not the case for the molecule HD, with one atom of deuterium (i.e., heavy hydrogen,  $^2\text{H}$ ). Because of its small moment of inertia, the HD molecule has a relatively large value of  $\epsilon$ : 0.0057 eV. At approximately what temperature would you expect the rotational heat capacity of a gas of HD molecules to “freeze out,” that is, to fall significantly below the constant value predicted by the equipartition theorem?

**Problem 6.30.** In this problem you will investigate the behavior of ordinary hydrogen,  $\text{H}_2$ , at low temperatures. The constant  $\epsilon$  is 0.0076 eV. As noted in the text, only half of the terms in the rotational partition function, equation 6.30, contribute for any given molecule. More precisely, the set of allowed  $j$  values is determined by the *spin* configuration of the two atomic nuclei. There are four independent spin configurations, classified as a single “singlet” state and three “triplet” states. The time required for a molecule to convert between the singlet and triplet configurations is ordinarily quite long, so the properties of the two types of molecules can be studied independently. The singlet molecules are known as **parahydrogen** while the triplet molecules are known as **orthohydrogen**.

- (a) For parahydrogen, only the rotational states with even values of  $j$  are allowed.\* Use a computer (as in Problem 6.28) to calculate the rotational

---

\*For those who have studied quantum mechanics, here's why: Even- $j$  wavefunctions are symmetric (unchanged) under the operation of replacing  $\vec{r}$  with  $-\vec{r}$ , which is equivalent to interchanging the two nuclei; odd- $j$  wavefunctions are antisymmetric under this operation. The two hydrogen nuclei (protons) are fermions, so their overall wavefunction must be antisymmetric under interchange. The singlet state ( $\uparrow\downarrow - \downarrow\uparrow$ ) is already antisymmetric in

partition function, average energy, and heat capacity of a parahydrogen molecule. Plot the heat capacity as a function of  $kT/\epsilon$ .\*

- (b) For orthohydrogen, only the rotational states with odd values of  $j$  are allowed. Repeat part (a) for orthohydrogen.
- (c) At high temperature, where the number of accessible even- $j$  states is essentially the same as the number of accessible odd- $j$  states, a sample of hydrogen gas will ordinarily consist of a mixture of 1/4 parahydrogen and 3/4 orthohydrogen. A mixture with these proportions is called **normal hydrogen**. Suppose that normal hydrogen is cooled to low temperature without allowing the spin configurations of the molecules to change. Plot the rotational heat capacity of this mixture as a function of temperature. At what temperature does the rotational heat capacity fall to half its high-temperature value (i.e., to  $k/2$  per molecule)?
- (d) Suppose now that some hydrogen is cooled in the presence of a catalyst that allows the nuclear spins to frequently change alignment. In this case *all* terms in the original partition function are allowed, but the odd- $j$  terms should be counted three times each because of the nuclear spin degeneracy. Calculate the rotational partition function, average energy, and heat capacity of this system, and plot the heat capacity as a function of  $kT/\epsilon$ .
- (e) A deuterium molecule,  $D_2$ , has nine independent nuclear spin configurations, of which six are “symmetric” and three are “antisymmetric.” The rule for nomenclature is that the variety with more independent states gets called “ortho-,” while the other gets called “para-.” For orthodeuterium only even- $j$  rotational states are allowed, while for paradeuterium only odd- $j$  states are allowed.<sup>†</sup> Suppose, then, that a sample of  $D_2$  gas, consisting of a normal equilibrium mixture of 2/3 ortho and 1/3 para, is cooled without allowing the nuclear spin configurations to change. Calculate and plot the rotational heat capacity of this system as a function of temperature.<sup>‡</sup>

### 6.3 The Equipartition Theorem

I’ve been invoking the equipartition theorem throughout this book, and we’ve verified that it is true in a number of particular cases, but so far I haven’t shown you an actual proof. The proof is quite easy, if you use Boltzmann factors.

The equipartition theorem doesn’t apply to all systems. It applies only to systems whose energy is in the form of quadratic “degrees of freedom,” of the form

$$E(q) = cq^2, \quad (6.34)$$

where  $c$  is a constant coefficient and  $q$  is any coordinate or momentum variable, like  $x$ , or  $p_x$ , or  $L_x$  (angular momentum). I’m going to treat just this single degree

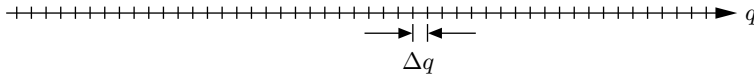
---

spin, so its spatial wavefunction must be symmetric, while the triplet states ( $\uparrow\uparrow$ ,  $\downarrow\downarrow$ , and  $\uparrow\downarrow + \downarrow\uparrow$ ) are symmetric in spin, so their spatial wavefunctions must be antisymmetric.

\*For a molecule such as  $O_2$  with spin-0 nuclei, this graph is the whole story; the only nuclear spin configuration is a singlet and only the even- $j$  states are allowed.

<sup>†</sup>Deuterium nuclei are bosons, so the overall wavefunction must be symmetric under interchange.

<sup>‡</sup>For a good discussion of hydrogen at low temperature, with references to experiments, see Gopal (1966).



**Figure 6.8.** To count states over a continuous variable  $q$ , pretend that they're discretely spaced, separated by  $\Delta q$ .

of freedom as my “system,” assume that it's in equilibrium with a reservoir at temperature  $T$ , and calculate its average energy,  $\overline{E}$ .

I'll analyze this system in classical mechanics, where each value of  $q$  corresponds to a separate, independent state. To *count* the states, I'll pretend that they're discretely spaced, separated by small intervals  $\Delta q$ , as shown in Figure 6.8. As long as  $\Delta q$  is extremely small, we expect it to cancel out of the final result for  $\overline{E}$ .

The partition function for this system is

$$Z = \sum_q e^{-\beta E(q)} = \sum_q e^{-\beta c q^2}. \quad (6.35)$$

To evaluate the sum, I'll multiply by  $\Delta q$  inside the sum and divide by  $\Delta q$  outside the sum:

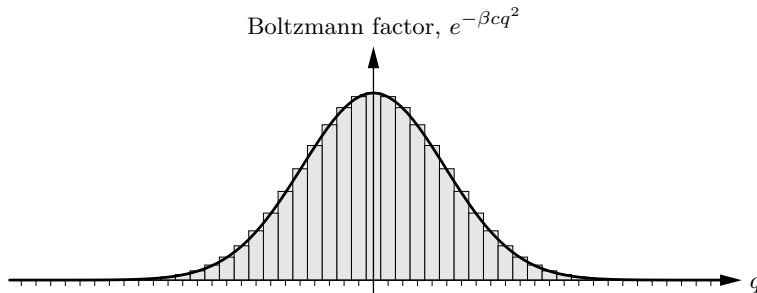
$$Z = \frac{1}{\Delta q} \sum_q e^{-\beta c q^2} \Delta q. \quad (6.36)$$

Now the sum can be interpreted as the area under a bar graph whose height is determined by the Boltzmann factor (see Figure 6.9). Since  $\Delta q$  is very small, we can approximate the bar graph by the smooth curve, changing the sum into an integral:

$$Z = \frac{1}{\Delta q} \int_{-\infty}^{\infty} e^{-\beta c q^2} dq. \quad (6.37)$$

Before trying to evaluate the integral, let's change variables to  $x = \sqrt{\beta c} q$ , so that  $dq = dx/\sqrt{\beta c}$ . Then

$$Z = \frac{1}{\Delta q} \frac{1}{\sqrt{\beta c}} \int_{-\infty}^{\infty} e^{-x^2} dx. \quad (6.38)$$



**Figure 6.9.** The partition function is the area under a bar graph whose height is the Boltzmann factor,  $e^{-\beta c q^2}$ . To calculate this area, we pretend that the bar graph is a smooth curve.

The integral over  $x$  is now just some number, whose value isn't very important as far as the physics is concerned. However, the integral is rather interesting mathematically. The function  $e^{-x^2}$  is called a **Gaussian**, and unfortunately its antiderivative cannot be written in terms of elementary functions. But there is a clever trick (described in Appendix B) for evaluating the *definite* integral from  $-\infty$  to  $\infty$ , and the result is simply  $\sqrt{\pi}$ . So our final result for the partition function is

$$Z = \frac{1}{\Delta q} \sqrt{\frac{\pi}{\beta c}} = C\beta^{-1/2}, \quad (6.39)$$

where  $C$  is just an abbreviation for  $\sqrt{\pi/c}/\Delta q$ .

Once you have an explicit formula for the partition function, it's easy to calculate the average energy, using the magical formula 6.25:

$$\begin{aligned} \bar{E} &= -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{1}{C\beta^{-1/2}} \frac{\partial}{\partial \beta} C\beta^{-1/2} \\ &= -\frac{1}{C\beta^{-1/2}} \left(-\frac{1}{2}\right) C\beta^{-3/2} = \frac{1}{2}\beta^{-1} = \frac{1}{2}kT. \end{aligned} \quad (6.40)$$

This is just the equipartition theorem. Notice that the constants  $c$ ,  $\Delta q$ , and  $\sqrt{\pi}$  have all canceled out.

The most important fact about this proof is that it does not carry over to quantum-mechanical systems. You can sort of see this from Figure 6.9: If the number of distinct states that have significant probabilities is too small, then the smooth Gaussian curve will not be a good approximation to the bar graph. And indeed, as we've seen in the case of an Einstein solid, the equipartition theorem is true only in the high-temperature limit, where many distinct states contribute and therefore the spacing between the states is unimportant. In general, the equipartition theorem applies only when the spacing between energy levels is much less than  $kT$ .

**Problem 6.31.** Consider a classical “degree of freedom” that is linear rather than quadratic:  $E = c|q|$  for some constant  $c$ . (An example would be the kinetic energy of a highly relativistic particle in one dimension, written in terms of its momentum.) Repeat the derivation of the equipartition theorem for this system, and show that the average energy is  $\bar{E} = kT$ .

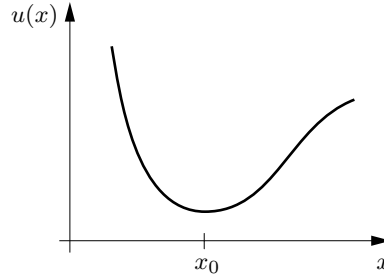
**Problem 6.32.** Consider a classical particle moving in a one-dimensional potential well  $u(x)$ , as shown in Figure 6.10. The particle is in thermal equilibrium with a reservoir at temperature  $T$ , so the probabilities of its various states are determined by Boltzmann statistics.

(a) Show that the average position of the particle is given by

$$\bar{x} = \frac{\int x e^{-\beta u(x)} dx}{\int e^{-\beta u(x)} dx},$$

where each integral is over the entire  $x$  axis.

**Figure 6.10.** A one-dimensional potential well. The higher the temperature, the farther the particle will stray from the equilibrium point.



- (b) If the temperature is reasonably low (but still high enough for classical mechanics to apply), the particle will spend most of its time near the bottom of the potential well. In that case we can expand  $u(x)$  in a Taylor series about the equilibrium point  $x_0$ :

$$u(x) = u(x_0) + (x - x_0) \left. \frac{du}{dx} \right|_{x_0} + \frac{1}{2} (x - x_0)^2 \left. \frac{d^2u}{dx^2} \right|_{x_0} + \frac{1}{3!} (x - x_0)^3 \left. \frac{d^3u}{dx^3} \right|_{x_0} + \cdots$$

Show that the linear term must be zero, and that truncating the series after the quadratic term results in the trivial prediction  $\bar{x} = x_0$ .

- (c) If we keep the cubic term in the Taylor series as well, the integrals in the formula for  $\bar{x}$  become difficult. To simplify them, assume that the cubic term is small, so its exponential can be expanded in a Taylor series (leaving the quadratic term in the exponent). Keeping only the largest temperature-dependent term, show that in this limit  $\bar{x}$  differs from  $x_0$  by a term proportional to  $kT$ . Express the coefficient of this term in terms of the coefficients of the Taylor series for  $u(x)$ .
- (d) The interaction of noble gas atoms can be modeled using the **Lennard-Jones potential**,

$$u(x) = u_0 \left[ \left( \frac{x_0}{x} \right)^{12} - 2 \left( \frac{x_0}{x} \right)^6 \right].$$

Sketch this function, and show that the minimum of the potential well is at  $x = x_0$ , with depth  $u_0$ . For argon,  $x_0 = 3.9 \text{ \AA}$  and  $u_0 = 0.010 \text{ eV}$ . Expand the Lennard-Jones potential in a Taylor series about the equilibrium point, and use the result of part (c) to predict the linear thermal expansion coefficient (see Problem 1.8) of a noble gas crystal in terms of  $u_0$ . Evaluate the result numerically for argon, and compare to the measured value  $\alpha = 0.0007 \text{ K}^{-1}$  (at 80 K).

## 6.4 The Maxwell Speed Distribution

For our next application of Boltzmann factors, I'd like to take a detailed look at the motion of molecules in an ideal gas. We already know (from the equipartition theorem) that the root-mean-square speed of the molecules is given by the formula

$$v_{\text{rms}} = \sqrt{\frac{3kT}{m}}. \quad (6.41)$$

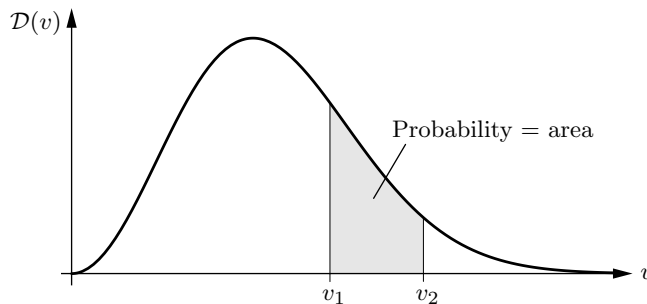
But this is just a sort of average. Some of the molecules will be moving faster than this, others slower. In practice, we might want to know exactly how many molecules are moving at any given speed. Equivalently, let's ask what is the probability of some particular molecule moving at a given speed.

Technically, the probability that a molecule is moving *at* any given speed  $v$  is *zero*. Since speed can vary continuously, there are infinitely many possible speeds, and therefore each of them has infinitesimal probability (which is essentially the same as zero). However, some speeds are less probable than others, and we can still represent the *relative* probabilities of various speeds by a graph, which turns out to look like Figure 6.11. The most probable speed is where the graph is the highest, and other speeds are less probable, in proportion to the height of the graph. Furthermore, if we normalize the graph (that is, adjust the vertical scale) in the right way, it has a more precise interpretation: The *area* under the graph between any two speeds  $v_1$  and  $v_2$  equals the probability that the molecule's speed is between  $v_1$  and  $v_2$ :

$$\text{Probability}(v_1 \dots v_2) = \int_{v_1}^{v_2} \mathcal{D}(v) dv, \quad (6.42)$$

where  $\mathcal{D}(v)$  is the height of the graph. If the interval between  $v_1$  and  $v_2$  is infinitesimal, then  $\mathcal{D}(v)$  doesn't change significantly within the interval and we can write simply

$$\text{Probability}(v \dots v+dv) = \mathcal{D}(v) dv. \quad (6.43)$$



**Figure 6.11.** A graph of the relative probabilities for a gas molecule to have various speeds. More precisely, the vertical scale is defined so that the area under the graph within any interval equals the probability of the molecule having a speed in that interval.

The function  $\mathcal{D}(v)$  is called a **distribution function**. Its actual *value* at any point isn't very meaningful by itself. Instead,  $\mathcal{D}(v)$  is a function *whose purpose in life is to be integrated*. To turn  $\mathcal{D}(v)$  into a probability you *must* integrate over some interval of  $v$ 's (or, if the interval is small, just multiply by the width of the interval). The function  $\mathcal{D}(v)$  itself doesn't even have the right *units* (namely, none) for a probability; instead, it has units of  $1/v$ , or  $(\text{m/s})^{-1}$ .

Now that we know how to interpret the answer, I'd like to derive a formula for the function  $\mathcal{D}(v)$ . The most important ingredient in the derivation is the Boltzmann factor. But another important element is the fact that space is three dimensional, which implies that for any given *speed*, there are many possible velocity *vectors*. In fact, we can write the function  $\mathcal{D}(v)$  schematically as

$$\mathcal{D}(v) \propto \left( \begin{array}{c} \text{probability of a molecule} \\ \text{having velocity } \vec{v} \end{array} \right) \times \left( \begin{array}{c} \text{number of vectors } \vec{v} \\ \text{corresponding to speed } v \end{array} \right). \quad (6.44)$$

There's also a constant of proportionality, which we'll worry about later.

The first factor in equation 6.44 is just the Boltzmann factor. Each velocity vector corresponds to a distinct molecular state, and the probability of a molecule being in any given state  $s$  is proportional to the Boltzmann factor  $e^{-E(s)/kT}$ . In this case the energy is just the translational kinetic energy,  $\frac{1}{2}mv^2$  (where  $v = |\vec{v}|$ ), so

$$\left( \begin{array}{c} \text{probability of a molecule} \\ \text{having velocity } \vec{v} \end{array} \right) \propto e^{-mv^2/2kT}. \quad (6.45)$$

I've neglected any variables besides velocity that might affect the state of the molecule, such as its position in space or its internal motion. This simplification is valid for an ideal gas, where the translational motion is independent of all other variables.

Equation 6.45 says that the most likely velocity *vector* for a molecule in an ideal gas is *zero*. Given what we know about Boltzmann factors, this result should hardly be surprising: Low-energy states are always more probable than high-energy states, for any system at finite (positive) temperature. However, the most likely velocity *vector* does *not* correspond to the most likely *speed*, because for some speeds there are *more* distinct velocity vectors than for others.

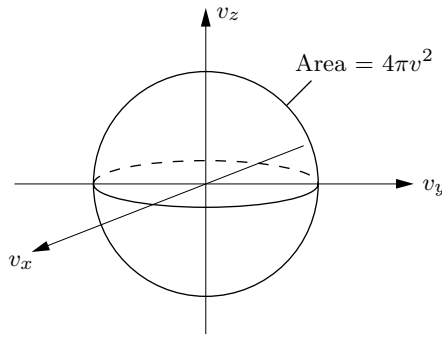
So let us turn to the second factor in equation 6.44. To evaluate this factor, imagine a three-dimensional "velocity space" in which each point represents a velocity vector (see Figure 6.12). The set of velocity vectors corresponding to any given speed  $v$  lives on the surface of a sphere with radius  $v$ . The larger  $v$  is, the bigger the sphere, and the more possible velocity vectors there are. So I claim that the second factor in equation 6.44 is the surface area of the sphere in velocity space:

$$\left( \begin{array}{c} \text{number of vectors } \vec{v} \\ \text{corresponding to speed } v \end{array} \right) \propto 4\pi v^2. \quad (6.46)$$

Putting this "degeneracy" factor together with the Boltzmann factor (6.45), we obtain

$$\mathcal{D}(v) = C \cdot 4\pi v^2 e^{-mv^2/2kT}, \quad (6.47)$$





**Figure 6.12.** In “velocity space” each point represents a possible velocity vector. The set of all vectors for a given speed  $v$  lies on the surface of a sphere with radius  $v$ .

where  $C$  is a constant of proportionality. To determine  $C$ , note that the total probability of finding the molecule at *some* speed must equal 1:

$$1 = \int_0^\infty \mathcal{D}(v) dv = 4\pi C \int_0^\infty v^2 e^{-mv^2/2kT} dv. \quad (6.48)$$

Changing variables to  $x = v\sqrt{m/2kT}$  puts this integral into the form

$$1 = 4\pi C \left( \frac{2kT}{m} \right)^{3/2} \int_0^\infty x^2 e^{-x^2} dx. \quad (6.49)$$

Like the pure Gaussian  $e^{-x^2}$ , the function  $x^2 e^{-x^2}$  cannot be anti-differentiated in terms of elementary functions. Again, however, there are tricks (explained in Appendix B) for evaluating the *definite* integral from 0 to  $\infty$ ; in this case the answer is  $\sqrt{\pi}/4$ . The 4 cancels, leaving us with  $C = (m/2\pi kT)^{3/2}$ .

Our final result for the distribution function  $\mathcal{D}(v)$  is therefore

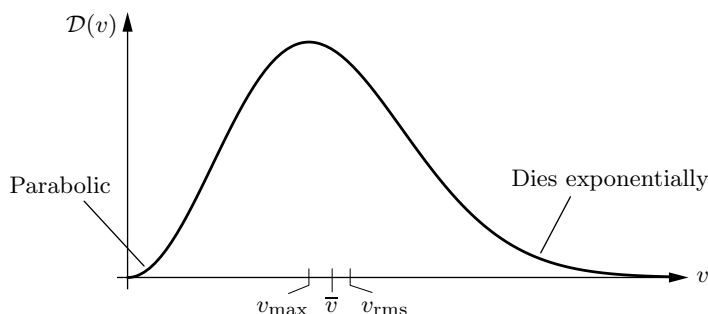
$$\mathcal{D}(v) = \left( \frac{m}{2\pi kT} \right)^{3/2} 4\pi v^2 e^{-mv^2/2kT}. \quad (6.50)$$

This result is called the **Maxwell distribution** (after James Clerk Maxwell) for the speeds of molecules in an ideal gas. It’s a complicated formula, but I hope you won’t find it hard to remember the important parts: the Boltzmann factor involving the translational kinetic energy, and the geometrical factor of the surface area of the sphere in velocity space.

Figure 6.13 shows another plot of the Maxwell distribution. At very small  $v$ , the Boltzmann factor is approximately 1 so the curve is a parabola; in particular, the distribution goes to zero at  $v = 0$ . This result does not contradict the fact that zero is the *most* likely velocity vector, because now we’re talking about *speeds*, and there are simply too few velocity vectors corresponding to very small speeds.

Meanwhile, the Maxwell distribution also goes to zero at very high speeds (much greater than  $\sqrt{kT/m}$ ), because of the exponential fall-off of the Boltzmann factor.

In between  $v = 0$  and  $v = \infty$ , the Maxwell distribution rises and falls. By setting the derivative of equation 6.50 equal to zero, you can show that the maximum value of  $\mathcal{D}(v)$  occurs at  $v_{\max} = \sqrt{2kT/m}$ . As you would expect, the peak shifts to the



**Figure 6.13.** The Maxwell speed distribution falls off as  $v \rightarrow 0$  and as  $v \rightarrow \infty$ . The average speed is slightly larger than the most likely speed, while the rms speed is a bit larger still.

right as the temperature is increased. The most likely speed is not the same as the rms speed; referring to equation 6.41, we see that the rms speed is the greater of the two, by about 22%. The *average* speed is different still; to compute it, add up all the possible speeds, weighted by their probabilities:

$$\bar{v} = \sum_{\text{all } v} v D(v) dv. \quad (6.51)$$

In this equation I'm imagining the various speeds to be discretely spaced, separated by  $dv$ . If you turn the sum into an integral and evaluate it, you get

$$\bar{v} = \sqrt{\frac{8kT}{\pi m}}, \quad (6.52)$$

which lies in between  $v_{\text{max}}$  and  $v_{\text{rms}}$ .

As an example, consider the nitrogen molecules in air at room temperature. You can easily calculate the most probable speed, which turns out to be 422 m/s at 300 K. But some of the molecules are moving much faster than this, while others are moving much slower. What is the probability that a particular molecule is moving faster than 1000 m/s?

First let's make a graphical estimate. The speed 1000 m/s exceeds  $v_{\text{max}}$  by a factor of

$$\frac{1000 \text{ m/s}}{422 \text{ m/s}} = 2.37. \quad (6.53)$$

Looking at Figure 6.13, you can see that at this point the Maxwell distribution is rapidly dying out but not yet dead. The area under the graph beyond  $2.37v_{\text{max}}$  looks like only one or two percent of the total area under the graph.

Quantitatively, the probability is given by the integral of the Maxwell distribution from 1000 m/s to infinity:

$$\text{Probability}(v > 1000 \text{ m/s}) = 4\pi \left( \frac{m}{2\pi kT} \right)^{3/2} \int_{1000 \text{ m/s}}^{\infty} v^2 e^{-mv^2/2kT} dv. \quad (6.54)$$

With this nontrivial lower limit, the integral *cannot* be carried out analytically; the best option is to do it numerically, by calculator or computer. You *could* go ahead and plug in numbers at this point, instructing the computer to work in units of m/s. But it's much cleaner to first change variables to  $x = v\sqrt{m/2kT} = v/v_{\max}$ , just as in equation 6.49. The integral then becomes

$$4\pi\left(\frac{m}{2\pi kT}\right)^{3/2}\left(\frac{2kT}{m}\right)^{3/2}\int_{x_{\min}}^{\infty}x^2e^{-x^2}dx = \frac{4}{\sqrt{\pi}}\int_{x_{\min}}^{\infty}x^2e^{-x^2}dx, \quad (6.55)$$

where the lower limit is the value of  $x$  when  $v = 1000$  m/s, that is,  $x_{\min} = (1000 \text{ m/s})/(422 \text{ m/s}) = 2.37$ . Now it's *easy* to type the integral into a computer. I did so and got an answer of 0.0105 for the probability. Only about 1% of the nitrogen molecules are moving faster than 1000 m/s.

**Problem 6.33.** Calculate the most probable speed, average speed, and rms speed for oxygen ( $\text{O}_2$ ) molecules at room temperature.

**Problem 6.34.** Carefully plot the Maxwell speed distribution for nitrogen molecules at  $T = 300$  K and at  $T = 600$  K. Plot both graphs on the same axes, and label the axes with numbers.

**Problem 6.35.** Verify from the Maxwell speed distribution that the most likely speed of a molecule is  $\sqrt{2kT/m}$ .

**Problem 6.36.** Fill in the steps between equations 6.51 and 6.52, to determine the average speed of the molecules in an ideal gas.

**Problem 6.37.** Use the Maxwell distribution to calculate the average value of  $v^2$  for the molecules in an ideal gas. Check that your answer agrees with equation 6.41.

**Problem 6.38.** At room temperature, what fraction of the nitrogen molecules in the air are moving at less than 300 m/s?

**Problem 6.39.** A particle near earth's surface traveling faster than about 11 km/s has enough kinetic energy to completely escape from the earth, despite earth's gravitational pull. Molecules in the upper atmosphere that are moving faster than this will therefore escape if they do not suffer any collisions on the way out.

- (a) The temperature of earth's upper atmosphere is actually quite high, around 1000 K. Calculate the probability of a nitrogen molecule at this temperature moving faster than 11 km/s, and comment on the result.
- (b) Repeat the calculation for a hydrogen molecule ( $\text{H}_2$ ) and for a helium atom, and discuss the implications.
- (c) Escape speed from the moon's surface is only about 2.4 km/s. Explain why the moon has no atmosphere.

**Problem 6.40.** You might wonder why all the molecules in a gas in thermal equilibrium don't have exactly the *same* speed. After all, when two molecules collide, doesn't the faster one always lose energy and the slower one gain energy? And if so, wouldn't repeated collisions eventually bring all the molecules to some common speed? Describe an example of a billiard-ball collision in which this is *not* the case: The faster ball *gains* energy and the slower ball *loses* energy. Include numbers, and be sure that your collision conserves both energy and momentum.

**Problem 6.41.** Imagine a world in which space is two-dimensional, but the laws of physics are otherwise the same. Derive the speed distribution formula for an ideal gas of nonrelativistic particles in this fictitious world, and sketch this distribution. Carefully explain the similarities and differences between the two-dimensional and three-dimensional cases. What is the most likely velocity vector? What is the most likely speed?

## 6.5 Partition Functions and Free Energy

For an isolated system with fixed energy  $U$ , the most fundamental statistical quantity is the multiplicity,  $\Omega(U)$ —the number of available microstates. The logarithm of the multiplicity gives the entropy, which tends to increase.

For a system in equilibrium with a reservoir at temperature  $T$  (see Figure 6.14), the quantity most analogous to  $\Omega$  is the partition function,  $Z(T)$ . Like  $\Omega(U)$ , the partition function is more or less equal to the number of microstates available to the system (but at fixed temperature, not fixed energy). We might therefore expect its logarithm to be a quantity that tends to increase under these conditions. But we already know a quantity that tends to *decrease* under these conditions: the Helmholtz free energy,  $F$ . The quantity that tends to *increase* would be  $-F$ , or, if we want a dimensionless quantity,  $-F/kT$ . Taking a giant intuitive leap, we might therefore guess the formula

$$F = -kT \ln Z \quad \text{or} \quad Z = e^{-F/kT}. \quad (6.56)$$

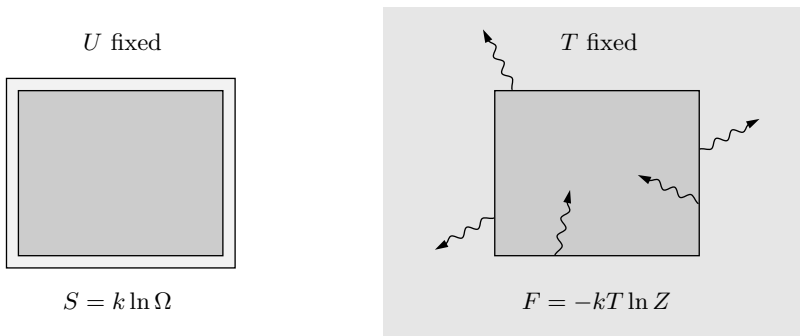
Indeed, this formula turns out to be true. Let me prove it.

First recall the definition of  $F$ :

$$F \equiv U - TS. \quad (6.57)$$

Also, recall the partial derivative relation

$$\left( \frac{\partial F}{\partial T} \right)_{V,N} = -S. \quad (6.58)$$



**Figure 6.14.** For an isolated system (left),  $S$  tends to increase. For a system at constant temperature (right),  $F$  tends to decrease. Like  $S$ ,  $F$  can be written as the logarithm of a statistical quantity, in this case  $Z$ .

Solving equation 6.57 for  $S$  and plugging into equation 6.58 gives

$$\left(\frac{\partial F}{\partial T}\right)_{V,N} = \frac{F - U}{T}. \quad (6.59)$$

This is a differential equation for the function  $F(T)$ , for any fixed  $V$  and  $N$ . To prove equation 6.56, I'll show that the quantity  $-kT \ln Z$  obeys the same differential equation, with the same “initial” condition at  $T = 0$ .

Let me define the symbol  $\tilde{F}$  to stand for the quantity  $-kT \ln Z$ . Holding  $V$  and  $N$  fixed, I want to find a formula for the derivative of this quantity with respect to  $T$ :

$$\frac{\partial \tilde{F}}{\partial T} = \frac{\partial}{\partial T}(-kT \ln Z) = -k \ln Z - kT \frac{\partial}{\partial T} \ln Z. \quad (6.60)$$

In the second term I'll use the chain rule to rewrite the derivative in terms of  $\beta = 1/kT$ :

$$\frac{\partial}{\partial T} \ln Z = \frac{\partial \beta}{\partial T} \frac{\partial}{\partial \beta} \ln Z = \frac{-1}{kT^2} \frac{1}{Z} \frac{\partial Z}{\partial \beta} = \frac{U}{kT^2}. \quad (6.61)$$

(Here I'm using  $U$  instead of  $\bar{E}$  for the system's average energy, because these ideas are most useful when applied to fairly large systems.) Plugging this result back into equation 6.60, we obtain

$$\frac{\partial \tilde{F}}{\partial T} = -k \ln Z - kT \frac{U}{kT^2} = \frac{\tilde{F}}{T} - \frac{U}{T}, \quad (6.62)$$

that is,  $\tilde{F}$  obeys exactly the same differential equation as  $F$ .

A first-order differential equation has an infinite family of solutions, corresponding to different “initial” conditions. So to complete the proof that  $\tilde{F} = F$ , I need to show that they're the same for at least one particular value of  $T$ , say  $T = 0$ . At  $T = 0$ , the original  $F$  is simply equal to  $U$ , the energy of the system when it is at zero temperature. This energy must be the lowest possible energy,  $U_0$ , since the Boltzmann factors  $e^{-U(s)/kT}$  for all excited states will be infinitely suppressed in comparison to the ground state. Meanwhile, the partition function at  $T = 0$  is simply  $e^{-U_0/kT}$ , again since all other Boltzmann factors are infinitely suppressed in comparison. Therefore

$$\tilde{F}(0) = -kT \ln Z(0) = U_0 = F(0), \quad (6.63)$$

completing the proof that  $\tilde{F} = F$  for all  $T$ .

The usefulness of the formula  $F = -kT \ln Z$  is that from  $F$  we can compute the entropy, pressure, and chemical potential, using the partial-derivative formulas

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N}, \quad P = -\left(\frac{\partial F}{\partial V}\right)_{T,N}, \quad \mu = +\left(\frac{\partial F}{\partial N}\right)_{T,V}. \quad (6.64)$$

In this way we can compute all the thermodynamic properties of a system, once we know its partition function. In Section 6.7 I'll apply this technique to analyze an ideal gas.

**Problem 6.42.** In Problem 6.20 you computed the partition function for a quantum harmonic oscillator:  $Z_{\text{h.o.}} = 1/(1 - e^{-\beta\epsilon})$ , where  $\epsilon = hf$  is the spacing between energy levels.

- (a) Find an expression for the Helmholtz free energy of a system of  $N$  harmonic oscillators.
- (b) Find an expression for the entropy of this system as a function of temperature. (Don't worry, the result is fairly complicated.)

**Problem 6.43.** Some advanced textbooks define entropy by the formula

$$S = -k \sum_s \mathcal{P}(s) \ln \mathcal{P}(s),$$

where the sum runs over all microstates accessible to the system and  $\mathcal{P}(s)$  is the probability of the system being in microstate  $s$ .

- (a) For an isolated system,  $\mathcal{P}(s) = 1/\Omega$  for all accessible states  $s$ . Show that in this case the preceding formula reduces to our familiar definition of entropy.
- (b) For a system in thermal equilibrium with a reservoir at temperature  $T$ ,  $\mathcal{P}(s) = e^{-E(s)/kT}/Z$ . Show that in this case as well, the preceding formula agrees with what we already know about entropy.

## 6.6 Partition Functions for Composite Systems

Before trying to write down the partition function for an ideal gas, it is useful to ask in general how the partition function for a system of several particles is related to the partition function for each individual particle. For instance, consider a system of just two particles, 1 and 2. If these particles do not interact with each other, so their total energy is simply  $E_1 + E_2$ , then

$$Z_{\text{total}} = \sum_s e^{-\beta[E_1(s) + E_2(s)]} = \sum_s e^{-\beta E_1(s)} e^{-\beta E_2(s)}, \quad (6.65)$$

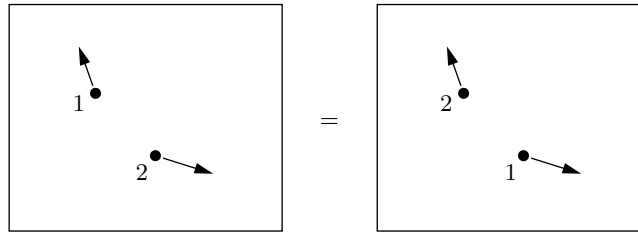
where the sum runs over all states  $s$  for the composite system. If, in addition, the two particles are distinguishable (either by their fixed positions or by some intrinsic properties), then the set of states for the composite system is equivalent to the set of all possible pairs of states,  $(s_1, s_2)$ , for the two particles individually. In this case,

$$Z_{\text{total}} = \sum_{s_1} \sum_{s_2} e^{-\beta E_1(s_1)} e^{-\beta E_2(s_2)}, \quad (6.66)$$

where  $s_1$  represents the state of particle 1 and  $s_2$  represents the state of particle 2. The first Boltzmann factor can be moved outside the sum over  $s_2$ . This sum, now of just the second Boltzmann factor, is simply the partition function for particle 2 alone,  $Z_2$ . This partition function is independent of  $s_1$ , and can therefore be taken out of the remaining sum. Finally, the sum over  $s_1$  gives simply  $Z_1$ , leaving us with

$$Z_{\text{total}} = Z_1 Z_2 \quad (\text{noninteracting, distinguishable particles}). \quad (6.67)$$

If the particles are *indistinguishable*, however, the step going from equation 6.65 to equation 6.66 is not valid. The problem is exactly the same as the one



**Figure 6.15.** Interchanging the states of two indistinguishable particles leaves the system in the same state as before.

we encountered when computing the multiplicity of an ideal gas in Section 2.5: Putting particle 1 in state  $A$  and particle 2 in state  $B$  is the same thing as putting particle 2 in state  $A$  and particle 1 in state  $B$  (see Figure 6.15). Equation 6.66 therefore counts nearly every state twice, and a more accurate formula would be

$$Z_{\text{total}} = \frac{1}{2} Z_1 Z_2 \quad (\text{noninteracting, indistinguishable particles}). \quad (6.68)$$

This formula still isn't precisely correct, because there are some terms in the double sum of equation 6.66 in which both particles are in the *same* state, that is,  $s_1 = s_2$ . These terms have *not* been double-counted, so we shouldn't divide their number by 2. But for ordinary gases and many other familiar systems, the density is low enough that the chances of both particles being in the same state are negligible. The terms with  $s_1 = s_2$  are therefore only a tiny fraction of all the terms in equation 6.66, and it doesn't much matter whether we count them correctly or not.\*

The generalization of equations 6.67 and 6.68 to systems of more than two particles is straightforward. If the particles are distinguishable, the total partition function is the product of all the individual partition functions:

$$Z_{\text{total}} = Z_1 Z_2 Z_3 \cdots Z_N \quad (\text{noninteracting, distinguishable systems}). \quad (6.69)$$

This equation also applies to the total partition function of a *single* particle that can store energy in several ways; for instance,  $Z_1$  could be the partition function for its motion in the  $x$  direction,  $Z_2$  for its motion in the  $y$  direction, and so on.

For a not-too-dense system of  $N$  indistinguishable particles, the general formula is

$$Z_{\text{total}} = \frac{1}{N!} Z_1^N \quad (\text{noninteracting, indistinguishable particles}), \quad (6.70)$$

where  $Z_1$  is the partition function for any one of the particles individually. The number of ways of interchanging  $N$  particles with each other is  $N!$ , hence the prefactor.

When we deal with multiparticle systems, one point of terminology can be confusing. It is important to distinguish the “state” of an individual particle from

---

\*The following chapter deals with very dense systems for which this issue is important. Until then, don't worry about it.

the “state” of the entire system. Unfortunately, I don’t know of a good concise way to distinguish between these two concepts. When the context is ambiguous, I’ll write **single-particle state** or **system state**, as appropriate. In the preceding discussion,  $s$  is the system state while  $s_1$  and  $s_2$  are single-particle states. In general, to specify the system state, you must specify the single-particle states of all the particles in the system.

**Problem 6.44.** Consider a large system of  $N$  indistinguishable, noninteracting molecules (perhaps in an ideal gas or a dilute solution). Find an expression for the Helmholtz free energy of this system, in terms of  $Z_1$ , the partition function for a single molecule. (Use Stirling’s approximation to eliminate the  $N!$ .) Then use your result to find the chemical potential, again in terms of  $Z_1$ .

## 6.7 Ideal Gas Revisited

### The Partition Function

We now have all the tools needed to calculate the partition function, and hence all the other thermal quantities, of an ideal gas. An *ideal* gas, as before, means one in which the molecules are usually far enough apart that we can neglect any energy due to forces between them. If the gas contains  $N$  molecules (all identical), then its partition function has the form

$$Z = \frac{1}{N!} Z_1^N, \quad (6.71)$$

where  $Z_1$  is the partition function for one individual molecule.

To calculate  $Z_1$ , we must add up the Boltzmann factors for all possible microstates of a single molecule. Each Boltzmann factor has the form

$$e^{-E(s)/kT} = e^{-E_{\text{tr}}(s)/kT} e^{-E_{\text{int}}(s)/kT}, \quad (6.72)$$

where  $E_{\text{tr}}$  is the molecule’s translational kinetic energy and  $E_{\text{int}}$  is its **internal energy** (rotational, vibrational, or whatever), for the state  $s$ . The sum over all single-particle states can be written as a double sum over translational states and internal states, allowing us to factor the partition function as in the previous section. The result is simply

$$Z_1 = Z_{\text{tr}} Z_{\text{int}}, \quad (6.73)$$

where

$$Z_{\text{tr}} = \sum_{\text{translational states}} e^{-E_{\text{tr}}/kT} \quad \text{and} \quad Z_{\text{int}} = \sum_{\text{internal states}} e^{-E_{\text{int}}/kT}. \quad (6.74)$$

The internal partition functions for rotational and vibrational states are treated in Section 6.2. For a given rotational and vibrational state, a molecule can also have various *electronic* states, in which its electrons are in different independent wavefunctions. For most molecules at ordinary temperatures, electronic *excited* states have negligible Boltzmann factors, due to their rather high energies. The electronic



*ground state*, however, can sometimes be degenerate. An oxygen molecule, for example, has a threefold-degenerate ground state, which contributes a factor of 3 to its internal partition function.

Now let us put aside the internal partition function and concentrate on the translational part,  $Z_{\text{tr}}$ . To compute  $Z_{\text{tr}}$ , we need to add up the Boltzmann factors for all possible translational states of a molecule. One way to enumerate these states is to count all the possible position and momentum vectors for a molecule, slipping in a factor of  $1/h^3$  to account for quantum mechanics as in Section 2.5. Instead, however, I'd now like to use the more rigorous method of counting all the independent definite-energy wavefunctions, just as we've been doing with internal states of atoms and molecules. I'll start with the case of a molecule confined to a one-dimensional box, then generalize to three dimensions.

A few of the definite-energy wavefunctions for a molecule in a one-dimensional box are shown in Figure 6.16. Because the molecule is confined to the box, its wavefunction must go to zero at each end, and therefore the allowed standing-wave patterns are limited to wavelengths of

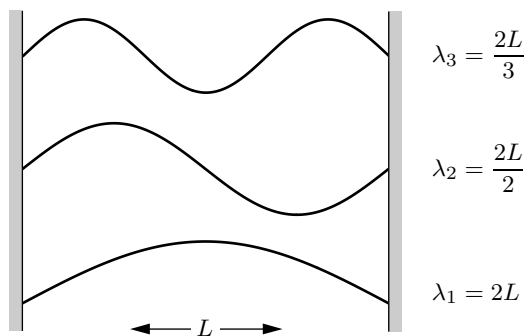
$$\lambda_n = \frac{2L}{n}, \quad n = 1, 2, \dots, \quad (6.75)$$

where  $L$  is the length of the box and  $n$  is the number of “bumps.” Each of these standing waves can be thought of as a superposition of left- and right-moving traveling waves with equal and opposite momenta; the magnitude of the momentum is given by the de Broglie relation  $p = h/\lambda$ , that is,

$$p_n = \frac{h}{\lambda_n} = \frac{hn}{2L}. \quad (6.76)$$

Finally, the relation between energy and momentum for a nonrelativistic particle is  $E = p^2/2m$ , where  $m$  is its mass. So the allowed energies for a molecule in a one-dimensional box are

$$E_n = \frac{p_n^2}{2m} = \frac{h^2 n^2}{8mL^2}. \quad (6.77)$$



**Figure 6.16.** The three lowest-energy wavefunctions for a particle confined to a one-dimensional box.

Knowing the energies, we can immediately write down the translational partition function for this molecule (still in one dimension):

$$Z_{1d} = \sum_n e^{-E_n/kT} = \sum_n e^{-h^2 n^2 / 8mL^2 kT}. \quad (6.78)$$

Unless  $L$  and/or  $T$  is *extremely* small, the energy levels are extremely close together, so we may as well approximate the sum as an integral:

$$Z_{1d} = \int_0^\infty e^{-h^2 n^2 / 8mL^2 kT} dn = \frac{\sqrt{\pi}}{2} \sqrt{\frac{8mL^2 kT}{h^2}} = \sqrt{\frac{2\pi mkT}{h^2}} L \equiv \frac{L}{\ell_Q}, \quad (6.79)$$

where  $\ell_Q$  is defined as the reciprocal of the square root in the previous expression:

$$\ell_Q \equiv \frac{h}{\sqrt{2\pi mkT}}. \quad (6.80)$$

I like to call  $\ell_Q$  the **quantum length**; aside from the factor of  $\pi$ , it is the de Broglie wavelength of a particle of mass  $m$  whose kinetic energy is  $kT$ . For a nitrogen molecule at room temperature, the quantum length works out to  $1.9 \times 10^{-11}$  m. The ratio  $L/\ell_Q$  is therefore quite large for any realistic box, meaning that *many* translational states are available to the molecule under these conditions: roughly the number of de Broglie wavelengths that would fit inside the box.

So much for a molecule moving in one dimension. For a molecule moving in three dimensions, the total kinetic energy is

$$E_{tr} = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m}, \quad (6.81)$$

where each momentum component can take on infinitely many different values according to formula 6.76. Since the  $n$ 's for the three momentum components can be chosen independently, we can again factor the partition function, into a piece for each of the three dimensions:

$$\begin{aligned} Z_{tr} &= \sum_s e^{-E_{tr}/kT} = \sum_{n_x} \sum_{n_y} \sum_{n_z} e^{-h^2 n_x^2 / 8mL_x^2 kT} e^{-h^2 n_y^2 / 8mL_y^2 kT} e^{-h^2 n_z^2 / 8mL_z^2 kT} \\ &= \frac{L_x}{\ell_Q} \frac{L_y}{\ell_Q} \frac{L_z}{\ell_Q} = \frac{V}{v_Q}, \end{aligned} \quad (6.82)$$

where  $V$  is the total volume of the box and  $v_Q$  is the **quantum volume**,

$$v_Q = \ell_Q^3 = \left( \frac{h}{\sqrt{2\pi mkT}} \right)^3. \quad (6.83)$$

The quantum volume is just the cube of the quantum length, so it's very small for a molecule at room temperature. The translational partition function is essentially the number of de Broglie-wavelength cubes that would fit inside the entire volume of the box, and is again quite large under ordinary conditions.

Combining this result with equation 6.73, we obtain for the single-particle partition function

$$Z_1 = \frac{V}{v_Q} Z_{\text{int}}, \quad (6.84)$$

where  $Z_{\text{int}}$  is a sum over all relevant internal states. The partition function for the entire gas of  $N$  molecules is then

$$Z = \frac{1}{N!} \left( \frac{V Z_{\text{int}}}{v_Q} \right)^N. \quad (6.85)$$

For future reference, the logarithm of the partition function is

$$\ln Z = N [\ln V + \ln Z_{\text{int}} - \ln N - \ln v_Q + 1]. \quad (6.86)$$

### Predictions

At this point we can compute all of the thermal properties of an ideal gas. Let's start with the total (average) energy, using the formula derived in Problem 6.16:

$$U = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial}{\partial \beta} \ln Z. \quad (6.87)$$

The quantities in equation 6.86 that depend on  $\beta$  are  $Z_{\text{int}}$  and  $v_Q$ , so

$$U = -N \frac{\partial}{\partial \beta} \ln Z_{\text{int}} + N \frac{1}{v_Q} \frac{\partial v_Q}{\partial \beta} = N \bar{E}_{\text{int}} + N \cdot \frac{3}{2} \frac{1}{\beta} = U_{\text{int}} + \frac{3}{2} NkT. \quad (6.88)$$

Here  $\bar{E}_{\text{int}}$  is the average internal energy of a molecule. The average translational kinetic energy is  $\frac{3}{2}kT$ , as we already knew from the equipartition theorem. Taking another derivative gives the heat capacity,

$$C_V = \frac{\partial U}{\partial T} = \frac{\partial U_{\text{int}}}{\partial T} + \frac{3}{2} Nk. \quad (6.89)$$

For a diatomic gas, the internal contribution to the heat capacity comes from rotation and vibration. As shown in Section 6.2, each of these contributions adds approximately  $Nk$  to the heat capacity at sufficiently high temperatures, but goes to zero at lower temperatures. The translational contribution could also freeze out in theory, but only at temperatures so low that  $\ell_Q$  is of order  $L$ , so replacing the sum by an integral in equation 6.79 becomes invalid. We have now explained all the features in the graph of  $C_V$  for hydrogen shown in Figure 1.13.

To compute the remaining thermal properties of an ideal gas, we need the Helmholtz free energy,

$$\begin{aligned} F = -kT \ln Z &= -NkT [\ln V + \ln Z_{\text{int}} - \ln N - \ln v_Q + 1] \\ &= -NkT [\ln V - \ln N - \ln v_Q + 1] + F_{\text{int}}, \end{aligned} \quad (6.90)$$

where  $F_{\text{int}}$  is the internal contribution to  $F$ , namely  $-NkT \ln Z_{\text{int}}$ . From this expression it's easy to compute the pressure,

$$P = -\left(\frac{\partial F}{\partial V}\right)_{T,N} = \frac{NkT}{V}. \quad (6.91)$$

I'll let you work out the entropy and chemical potential. The results are

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} = Nk \left[ \ln \left( \frac{V}{Nv_Q} \right) + \frac{5}{2} \right] - \frac{\partial F_{\text{int}}}{\partial T} \quad (6.92)$$

and

$$\mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} = -kT \ln \left( \frac{V Z_{\text{int}}}{Nv_Q} \right). \quad (6.93)$$

If we neglect the internal contributions, both of these quantities reduce to our earlier results for a monatomic ideal gas.

**Problem 6.45.** Derive equations 6.92 and 6.93 for the entropy and chemical potential of an ideal gas.

**Problem 6.46.** Equations 6.92 and 6.93 for the entropy and chemical potential involve the logarithm of the quantity  $VZ_{\text{int}}/Nv_Q$ . Is this logarithm normally positive or negative? Plug in some numbers for an ordinary gas and discuss.

**Problem 6.47.** Estimate the temperature at which the translational motion of a nitrogen molecule would freeze out, in a box of width 1 cm.

**Problem 6.48.** For a diatomic gas near room temperature, the internal partition function is simply the rotational partition function computed in Section 6.2, multiplied by the degeneracy  $Z_e$  of the electronic ground state.

(a) Show that the entropy in this case is

$$S = Nk \left[ \ln \left( \frac{V Z_e Z_{\text{rot}}}{Nv_Q} \right) + \frac{7}{2} \right].$$

Calculate the entropy of a mole of oxygen ( $Z_e = 3$ ) at room temperature and atmospheric pressure, and compare to the measured value in the table at the back of this book.\*

(b) Calculate the chemical potential of oxygen in earth's atmosphere near sea level, at room temperature. Express the answer in electron-volts.

**Problem 6.49.** For a mole of nitrogen ( $\text{N}_2$ ) gas at room temperature and atmospheric pressure, compute the following:  $U$ ,  $H$ ,  $F$ ,  $G$ ,  $S$ , and  $\mu$ . (The rotational constant  $\epsilon$  for  $\text{N}_2$  is 0.00025 eV. The electronic ground state is not degenerate.)

**Problem 6.50.** Show explicitly from the results of this section that  $G = N\mu$  for an ideal gas.

---

\*See Rock (1983) or Gopal (1966) for a discussion of the comparison of theoretical and experimental entropies.

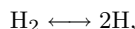
**Problem 6.51.** In this section we computed the single-particle translational partition function,  $Z_{\text{tr}}$ , by summing over all definite-energy wavefunctions. An alternative approach, however, is to sum over all possible position and momentum vectors, as we did in Section 2.5. Because position and momentum are continuous variables, the sums are really integrals, and we need to slip in a factor of  $1/h^3$  to get a unitless number that actually counts the independent wavefunctions. Thus, we might guess the formula

$$Z_{\text{tr}} = \frac{1}{h^3} \int d^3r d^3p e^{-E_{\text{tr}}/kT},$$

where the single integral sign actually represents six integrals, three over the position components (denoted  $d^3r$ ) and three over the momentum components (denoted  $d^3p$ ). The region of integration includes all momentum vectors, but only those position vectors that lie within a box of volume  $V$ . By evaluating the integrals explicitly, show that this expression yields the same result for the translational partition function as that obtained in the text. (The only time this formula would not be valid would be when the box is so small that we could not justify converting the sum in equation 6.78 to an integral.)

**Problem 6.52.** Consider an ideal gas of highly relativistic particles (such as photons or fast-moving electrons), whose energy-momentum relation is  $E = pc$  instead of  $E = p^2/2m$ . Assume that these particles live in a one-dimensional universe. By following the same logic that led to equation 6.79, derive a formula for the single-particle partition function,  $Z_1$ , for one particle in this gas.

**Problem 6.53.** The dissociation of molecular hydrogen into atomic hydrogen,



can be treated as an ideal gas reaction using the techniques of Section 5.6. The equilibrium constant  $K$  for this reaction is defined as

$$K = \frac{P_{\text{H}}^2}{P^\circ P_{\text{H}_2}},$$

where  $P^\circ$  is a reference pressure conventionally taken to be 1 bar, and the other  $P$ 's are the partial pressures of the two species at equilibrium. Now, using the methods of Boltzmann statistics developed in this chapter, you are ready to calculate  $K$  from first principles. Do so. That is, derive a formula for  $K$  in terms of more basic quantities such as the energy needed to dissociate one molecule (see Problem 1.53) and the internal partition function for molecular hydrogen. This internal partition function is a product of rotational and vibrational contributions, which you can estimate using the methods and data in Section 6.2. (An  $\text{H}_2$  molecule doesn't have any electronic spin degeneracy, but an H atom does—the electron can be in two different spin states. Neglect electronic excited states, which are important only at very high temperatures. The degeneracy due to nuclear spin alignments cancels, but include it if you wish.) Calculate  $K$  numerically at  $T = 300$  K, 1000 K, 3000 K, and 6000 K. Discuss the implications, working out a couple of numerical examples to show when hydrogen is mostly dissociated and when it is not.