2 The Second Law

The previous chapter explored the law of energy conservation as it applies to thermodynamic systems. It also introduced the concepts of heat, work, and temperature. However, some very fundamental questions remain unanswered: What is temperature, *really*, and why does heat flow spontaneously from a hotter object to a cooler object, never the other way? More generally, why do so many thermodynamic processes happen in one direction but never the reverse? This is the Big Question of thermal physics, which we now set out to answer.

In brief, the answer is this: Irreversible processes are not *inevitable*, they are just overwhelmingly *probable*. For instance, when heat flows from a hot object to a cooler object, the energy is just moving around more or less randomly. After we wait a while, the chances are overwhelming that we will find the energy distributed more "uniformly" (in a sense that I will make precise later) among all the parts of a system. "Temperature" is a way of quantifying the tendency of energy to enter or leave an object during the course of these random rearrangements.

To make these ideas precise, we need to study *how* systems store energy, and learn to count all the ways that the energy might be arranged. The mathematics of counting ways of arranging things is called **combinatorics**, and this chapter begins with a brief introduction to this subject.

2.1 Two-State Systems

Suppose that I flip three coins: a penny, a nickel, and a dime. How many possible outcomes are there? Not very many, so I've listed them all explicitly in Table 2.1. By this brute-force method, I count *eight* possible outcomes. If the coins are fair, each outcome is equally probable, so the probability of getting three heads or three tails is one in eight. There are three different ways of getting two heads and a tail, so the probability of getting exactly two heads is 3/8, as is the probability of

Penny	Nickel	Dime
H	H	H
H	H	T
H	T	H
T	H	H
H	T	T
T	H	T
T	T	H
T	T	T

Table 2.1. A list of all possible "microstates" of a set of three coins (where H is for heads and T is for tails).

getting exactly one head and two tails.

Now let me introduce some fancy terminology. Each of the eight different outcomes is called a **microstate**. In general, to specify the microstate of a system, we must specify the state of each individual particle, in this case the state of each coin. If we specify the state more generally, by merely saying how *many* heads or tails there are, we call it a **macrostate**. Of course, if you know the microstate of the system (say HHT), then you also know its macrostate (in this case, two heads). But the reverse is not true: Knowing that there are exactly two heads does *not* tell you the state of each coin, since there are three microstates corresponding to this macrostate. The number of microstates corresponding to a given macrostate is called the **multiplicity** of that macrostate, in this case 3.

The symbol I'll use for multiplicity is the Greek letter capital omega, Ω. In the example of the three coins, $\Omega(3 \text{ heads}) = 1$, $\Omega(2 \text{ heads}) = 3$, $\Omega(1 \text{ head}) = 3$, and $\Omega(0 \text{ heads}) = 1$. Note that the total multiplicity of all four macrostates is $1 + 3 + 3 + 1 = 8$, the total number of microstates. I'll call this quantity $\Omega(\text{all})$. Then the *probability* of any particular macrostate can be written

$$\text{probability of } n \text{ heads} = \frac{\Omega(n)}{\Omega(\text{all})}. \tag{2.1}$$

For instance, the probability of getting 2 heads is $\Omega(2)/\Omega(\text{all}) = 3/8$. Again, I'm assuming here that the coins are fair, so that all 8 microstates are equally probable.

To make things a little more interesting, suppose now that there are not just three coins but 100. The total number of *micro*states is now very large: 2^{100}, since each of the 100 coins has two possible states. The number of *macro*states, however, is only 101: 0 heads, 1 head, ... up to 100 heads. What about the multiplicities of these macrostates?

Let's start with the 0-heads macrostate. If there are zero heads, then every coin faces tails-up, so the exact microstate has been specified, that is, $\Omega(0) = 1$.

What if there is exactly one head? Well, the heads-up coin could be the first one, or the second one, etc., up to the 100th one; that is, there are exactly 100 possible microstates: $\Omega(1) = 100$. If you imagine all the coins starting heads-down, then $\Omega(1)$ is the number of ways of *choosing* one of them to turn over.

To find $\Omega(2)$, consider the number of ways of choosing two coins to turn heads-up. You have 100 choices for the first coin, and for each of these choices you have

99 remaining choices for the second coin. But you could choose any pair in either order, so the number of *distinct* pairs is

$$\Omega(2) = \frac{100 \cdot 99}{2}. \tag{2.2}$$

If you're going to turn three coins heads-up, you have 100 choices for the first, 99 for the second, and 98 for the third. But any triplet could be chosen in several ways: 3 choices for which one to flip first, and for each of these, 2 choices for which to flip second. Thus, the number of distinct triplets is

$$\Omega(3) = \frac{100 \cdot 99 \cdot 98}{3 \cdot 2}. \tag{2.3}$$

Perhaps you can now see the pattern. To find $\Omega(n)$, we write the product of n factors, starting with 100 and counting down, in the numerator. Then we divide by the product of n factors, starting with n and counting down to 1:

$$\Omega(n) = \frac{100 \cdot 99 \cdots (100 - n + 1)}{n \cdots 2 \cdot 1}. \tag{2.4}$$

The denominator is just n-factorial, denoted "$n!$". We can also write the numerator in terms of factorials, as $100!/(100-n)!$. (Imagine writing the product of all integers from 100 down to 1, then canceling all but the first n of them.) Thus the general formula can be written

$$\Omega(n) = \frac{100!}{n! \cdot (100-n)!} \equiv \binom{100}{n}. \tag{2.5}$$

The last expression is just a standard abbreviation for this quantity, sometimes spoken "100 choose n"—the number of different ways of choosing n items out of 100, or the number of "combinations" of n items chosen from 100.

If instead there are N coins, the multiplicity of the macrostate with n heads is

$$\Omega(N, n) = \frac{N!}{n! \cdot (N-n)!} = \binom{N}{n}, \tag{2.6}$$

the number of ways of choosing n objects out of N.

Problem 2.1. Suppose you flip four fair coins.

- **(a)** Make a list of all the possible outcomes, as in Table 2.1.
- **(b)** Make a list of all the different "macrostates" and their probabilities.
- **(c)** Compute the multiplicity of each macrostate using the combinatorial formula 2.6, and check that these results agree with what you got by brute-force counting.

Problem 2.2. Suppose you flip 20 fair coins.

- **(a)** How many possible outcomes (microstates) are there?
- **(b)** What is the probability of getting the sequence HTHHTTTHTHHHTHHHHTHT (in exactly that order)?
- **(c)** What is the probability of getting 12 heads and 8 tails (in any order)?

Problem 2.3. Suppose you flip 50 fair coins.

(a) How many possible outcomes (microstates) are there?
(b) How many ways are there of getting exactly 25 heads and 25 tails?
(c) What is the probability of getting exactly 25 heads and 25 tails?
(d) What is the probability of getting exactly 30 heads and 20 tails?
(e) What is the probability of getting exactly 40 heads and 10 tails?
(f) What is the probability of getting 50 heads and no tails?
(g) Plot a graph of the probability of getting n heads, as a function of n.

Problem 2.4. Calculate the number of possible five-card poker hands, dealt from a deck of 52 cards. (The order of cards in a hand does not matter.) A royal flush consists of the five highest-ranking cards (ace, king, queen, jack, 10) of any one of the four suits. What is the probability of being dealt a royal flush (on the first deal)?

The Two-State Paramagnet

You may be wondering what this silly coin-flipping example has to do with physics. Not much yet, but actually there are important physical systems whose combinatorics are exactly the same. Perhaps the most important of these is a **two-state paramagnet**.

All materials will respond in some way to a magnetic field, because of the electrical nature of electrons and atomic nuclei. A **paramagnet** is a material in which the constituent particles act like tiny compass needles that tend to align *parallel* to any externally applied magnetic field. (If the particles interact strongly enough with each other, the material can magnetize even *without* any externally applied field. We then call it a **ferromagnet**, after the most famous example, iron. Paramagnetism, in constrast, is a magnetic alignment that lasts only as long as an external field is applied.)

I'll refer to the individual magnetic particles as **dipoles**, because each has its own magnetic dipole moment vector. In practice each dipole could be an individual electron, a group of electrons in an atom, or an atomic nucleus. For any such microscopic dipole, quantum mechanics allows the component of the dipole moment vector along any given axis can take on only certain discrete values—intermediate values are not allowed. In the simplest case only *two* values are allowed, one positive and the other negative. We then have a **two-state paramagnet**, in which each elementary compass needle can have only two possible orientations, either parallel or antiparallel to the applied field. I'll draw this system as a bunch of little arrows, each pointing either up or down, as in Figure 2.1.[*]

Now for the combinatorics. Let's define $N_\uparrow$ to be the number of elementary dipoles that point up (at some particular time), and $N_\downarrow$ to be the number of dipoles that point down. The total number of dipoles is then $N = N_\uparrow + N_\downarrow$, and we'll

---

[*] A particle's dipole moment vector is proportional to its angular momentum vector; the simple two-state case occurs for particles with "spin 1/2." For a more complete discussion of quantum mechanics and angular momentum, see Appendix A.

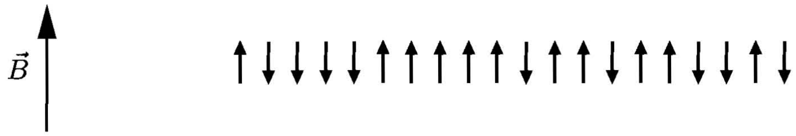


Figure 2.1. A symbolic representation of a two-state paramagnet, in which each elementary dipole can point either parallel or antiparallel to the externally applied magnetic field.

consider this number to be fixed. This system has one macrostate for each possible value of $N_\uparrow$, from 0 to N. The multiplicity of any macrostate is given by the same formula as in the coin-tossing example:

$$\Omega(N_\uparrow) = \binom{N}{N_\uparrow} = \frac{N!}{N_\uparrow!\,N_\downarrow!}. \tag{2.7}$$

The external magnetic field exerts a torque on each little dipole, trying to twist it to point parallel to the field. If the external field points up, then an up-dipole has *less* energy than a down-dipole, since you would have add energy to twist it from up to down. The total energy of the system (neglecting any interactions *between* dipoles) is determined by the total numbers of up- and down-dipoles, so specifying which macrostate this system is in is the same as specifying its total energy. In fact, in nearly all physical examples, the macrostate of a system is characterized, at least in part, by its total energy.

2.2 The Einstein Model of a Solid

Now let's move on to a system that's a bit more complicated, but also more representative of the systems typically encountered in physics. Consider a collection of microscopic systems that can each store any number of energy "units," all of the same size. Equal-size energy units occur for any quantum-mechanical *harmonic oscillator*, whose potential energy function has the form $\frac{1}{2}k_s x^2$ (where k_s is the "spring constant"). The size of the energy units is then hf,* where h is **Planck's constant** (6.63×10^{-34} J·s) and f is the natural frequency of the oscillator ($\frac{1}{2\pi}\sqrt{k_s/m}$). An abstract way of picturing a collection of many such oscillators is shown in Figure 2.2.

---

*As explained in Appendix A, the lowest possible energy of a quantum harmonic oscillator is actually $\frac{1}{2}hf$, not zero. But this "zero-point" energy never moves around, so it plays no role in thermal interactions. The excited-state energies are $\frac{3}{2}hf$, $\frac{5}{2}hf$, and so on, each with an additional energy "unit" of hf. For our purposes, it's fine to measure all energies relative to the ground state; then the allowed energies are 0, hf, $2hf$, etc.

Elsewhere you may see the energy unit of a quantum oscillator written as $\hbar\omega$, where $\hbar = h/2\pi$ and $\omega = 2\pi f$. The difference between $\hbar\omega$ and hf is nothing but a matter of where to put the factors of 2π.

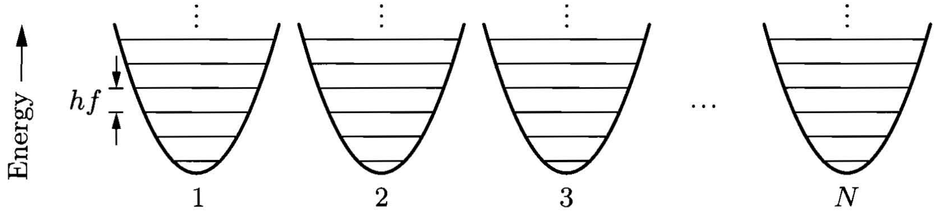


Figure 2.2. In quantum mechanics, any system with a quadratic potential energy function has evenly spaced energy levels separated in energy by hf, where f is the classical oscillation frequency. An Einstein solid is a collection of N such oscillators, all with the same frequency.

Examples of quantum oscillators include the vibrational motions of diatomic and polyatomic gas molecules. But an even more common example is the oscillation of atoms in a solid (see Figure 1.6). In a three-dimensional solid, each atom can oscillate in three independent directions, so if there are N oscillators, there are only $N/3$ atoms. The model of a solid as a collection of identical oscillators with quantized energy units was first proposed by Albert Einstein in 1907, so I will refer to this system as an **Einstein solid**.

Let's start with a very small Einstein solid, containing only three oscillators: $N = 3$. Table 2.2 lists the various microstates that this system could have, in order of increasing total energy; each row in the table corresponds to a different microstate. There is just one microstate with total energy 0, while there are three microstates with one unit of energy, six with two units, and ten with three units. That is,

$$\Omega(0) = 1, \qquad \Omega(1) = 3, \qquad \Omega(2) = 6, \qquad \Omega(3) = 10. \tag{2.8}$$

The general formula for the multiplicity of an Einstein solid with N oscillators

Oscillator:	#1	#2	#3
Energy:	0	0	0
	1	0	0
	0	1	0
	0	0	1
	2	0	0
	0	2	0
	0	0	2
	1	1	0
	1	0	1
	0	1	1

Oscillator:	#1	#2	#3
Energy:	3	0	0
	0	3	0
	0	0	3
	2	1	0
	2	0	1
	1	2	0
	0	2	1
	1	0	2
	0	1	2
	1	1	1

Table 2.2. Microstates of a small Einstein solid consisting of only three oscillators, containing a total of zero, one, two, or three units of energy.

and q energy units is

$$\Omega(N,q) = \binom{q+N-1}{q} = \frac{(q+N-1)!}{q!\,(N-1)!}. \tag{2.9}$$

Please check this formula for the examples just given. To *prove* this formula, let me adopt the following graphical representation of the microstate of an Einstein solid: I'll use a dot to represent each energy unit, and a vertical line to represent a partition between one oscillator and the next. So in a solid with four oscillators, the sequence

$$\bullet \mid \bullet\,\bullet\,\bullet \mid \; \mid \bullet\,\bullet\,\bullet\,\bullet$$

represents the microstate in which the first oscillator has one unit of energy, the second oscillator has three, the third oscillator has none, and the fourth oscillator has four. Notice that any microstate can be represented uniquely in this way, and that every possible sequence of dots and lines corresponds to a microstate. There are always q dots and $N-1$ lines, for a total of $q+N-1$ symbols. Given q and N, the number of possible arrangements is just the number of ways of choosing q of the symbols to be dots, that is, $\binom{q+N-1}{q}$.

Problem 2.5. For an Einstein solid with each of the following values of N and q, list all of the possible microstates, count them, and verify formula 2.9.

(a) $N = 3$, $q = 4$
(b) $N = 3$, $q = 5$
(c) $N = 3$, $q = 6$
(d) $N = 4$, $q = 2$
(e) $N = 4$, $q = 3$
(f) $N = 1$, $q =$ anything
(g) $N =$ anything, $q = 1$

Problem 2.6. Calculate the multiplicity of an Einstein solid with 30 oscillators and 30 units of energy. (Do not attempt to list all the microstates.)

Problem 2.7. For an Einstein solid with four oscillators and two units of energy, represent each possible microstate as a series of dots and vertical lines, as used in the text to prove equation 2.9.

> *You know, the most amazing thing happened to me tonight. I was coming here, on the way to the lecture, and I came in through the parking lot. And you won't believe what happened. I saw a car with the license plate ARW 357! Can you imagine? Of all the millions of license plates in the state, what was the chance that I would see that particular one tonight? Amazing!*
>
> —Richard Feynman, quoted by David Goodstein, *Physics Today* **42**, 73 (February, 1989).

2.3 Interacting Systems

We now know how to count the microstates of an Einstein solid. To understand heat flow and irreversible processes, however, we need to consider a system of *two* Einstein solids that can share energy back and forth.* I'll call the two solids A and B (see Figure 2.3).

First I should be clear about what is meant by the "macrostate" of such a composite system. For simplicity, I'll assume that the two solids are **weakly coupled**, so that the exchange of energy between them is much slower than the exchange of energy among atoms within each solid. Then the individual energies of the solids, U_A and U_B, will change only slowly; over sufficiently short time scales they are essentially fixed. I will use the word "macrostate" to refer to the state of the combined system, as specified by the (temporarily) constrained values of U_A and U_B. For any such macrostate we can compute the multiplicity, as we shall soon see. However, on longer time scales the values of U_A and U_B *will* change, so I'll also talk about the total multiplicity for *all* allowed values of U_A and U_B, counting all possible microstates with only the sum $U_{\text{total}} = U_A + U_B$ held fixed.

Let's start with a very small system, in which each of the "solids" contains only three harmonic oscillators and they contain a total of six units of energy:

$$N_A = N_B = 3; \qquad q_{\text{total}} = q_A + q_B = 6. \tag{2.10}$$

(Again I'm using q to denote the number of units of energy. The actual value of the energy is $U = qhf$.) Given these parameters, I must still specify the individual value of q_A or q_B to describe the macrostate of the system. There are seven possible macrostates, with $q_A = 0, 1, \ldots, 6$, as listed in Figure 2.4. I've used the standard formula $\binom{q+N-1}{q}$ to compute the individual multiplicities Ω_A and Ω_B for each macrostate. (I also computed some of them in the previous section by explicitly counting the microstates.) The total multiplicity of any macrostate, Ω_{total}, is just the product of the individual multiplicities, since the systems are independent of each other: For *each* of the Ω_A microstates available to solid A, there are Ω_B microstates available to solid B. The total multiplicity is also plotted in the bar

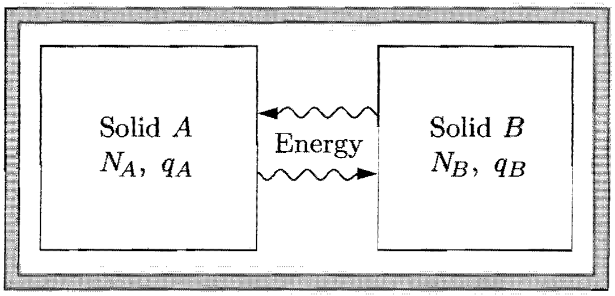


Figure 2.3. Two Einstein solids that can exchange energy with each other, isolated from the rest of the universe.

---

*This section and parts of Sections 3.1 and 3.3 are based on an article by T. A. Moore and D. V. Schroeder, *American Journal of Physics* **65**, 26–36 (1997).

q_A	Ω_A	q_B	Ω_B	$\Omega_{\text{total}} = \Omega_A\Omega_B$
0	1	6	28	28
1	3	5	21	63
2	6	4	15	90
3	10	3	10	100
4	15	2	6	90
5	21	1	3	63
6	28	0	1	28
				$462 = \binom{6+6-1}{6}$

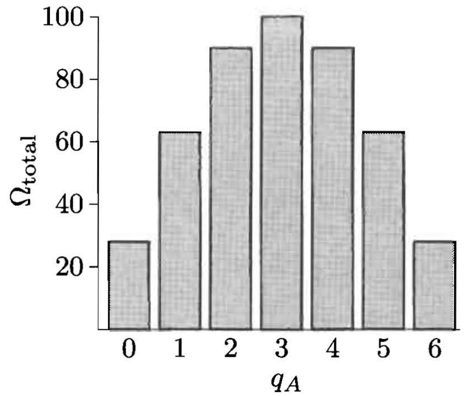


Figure 2.4. Macrostates and multiplicities of a system of two Einstein solids, each containing three oscillators, sharing a total of six units of energy.

graph. Over long time scales, the number of microstates accessible to the system is 462, the sum of the last column in the table. This number can also be checked by applying the standard formula to the entire system of six oscillators and six energy units.

Now let me introduce a big assumption: Let's assume that, over long time scales, the energy gets passed around randomly* in such a way that *all 462 microstates are equally probable.* So if you look at the system at any instant, you are equally likely to find it in any of the 462 microstates. This assumption is called the **fundamental assumption of statistical mechanics**:

> In an isolated system in thermal equilibrium, all accessible microstates are equally probable.

I can't prove this assumption, though it should seem plausible. At the microscopic level, we expect that any process that would take the system from state X to state Y is reversible, so that the system can just as easily go from state Y to state X.† In that case, the system should have no preference for one state over another. Still, it's not obvious that all of the supposedly "accessible" microstates can actually be reached within a reasonable amount of time. In fact, we'll soon see that for a large system, the number of "accessible" microstates is usually so huge that only a miniscule fraction of them could possibly occur within a lifetime. What we're assuming is that the microstates that *do* occur, over "long" but not unthinkably long time scales, constitute a representative sample. We assume that the transitions

---

*Exchange of energy requires some kind of interaction among the oscillators. Fortunately, the precise nature of this interaction doesn't really matter. There is a danger, though, that interactions among oscillators could affect the energy levels of each particular oscillator. This would spoil our assumption that the energy levels of each oscillator are evenly spaced. Let us therefore assume that the interactions among oscillators are strong enough to allow the exchange of energy, but too weak to have much effect on the energy levels themselves. This assumption is not fundamental to statistical mechanics, but it makes explicit calculations a whole lot easier.

†This idea is called the **principle of detailed balance**.

are "random," in the sense that they have no pattern that we could possibly care about.*

If we invoke the fundamental assumption for our system of two small Einstein solids, we can immediately conclude that, while all 462 *micro*states are equally probable, some *macro*states are more probable than others. The chance of finding the system in the fourth macrostate (with three energy units in each solid) is 100/462, while the chance of finding it in the first macrostate (with all the energy in solid B) is only 28/462. If all the energy is in solid B initially, and we wait a while, chances are we'll find the energy distributed more evenly later on.

Even for this very small system of only a few oscillators and energy units, computing all the multiplicities by hand is a bit of a chore. I would hate to do it for a system of a hundred oscillators and energy units. Fortunately, it's not hard to instruct a *computer* to do the arithmetic. Using a computer spreadsheet program, or comparable software, or perhaps even a graphing calculator, you should be able to reproduce the table and graph in Figure 2.4 without too much difficulty (see Problem 2.9).

Figure 2.5 shows a computer-generated table and graph for a system of two Einstein solids with

$$N_A = 300, \qquad N_B = 200, \qquad q_{\text{total}} = 100. \tag{2.11}$$

Now there are 101 possible macrostates, of which only a few are shown in the table. Look at the multiplicities: Even the *least* likely macrostate, with all the energy in solid B, has a multiplicity of 3×10^{81}. The *most* likely macrostate, with $q_A = 60$, has a multiplicity of 7×10^{114}. But what is important about these numbers is not that they are large, but that their *ratio* is large: The most likely macrostate is more than 10^{33} times more probable than the least likely macrostate.

Let's look at this example in a little more detail. The total number of microstates for all the macrostates is 9×10^{115}, so the probability of finding the system in its *most* likely macrostate is not particularly large: about 7%. There are several other macrostates, with q_A slightly smaller or larger than 60, whose probabilities are nearly as large. But as q_A gets farther away from 60, on either side, the probability drops off very sharply. The probability of finding q_A to be less than 30 or greater than 90 is less than one in a million, and the probability of finding $q_A < 10$ is less than 10^{-20}. The age of the universe is less than 10^{18} seconds, so you would need to check this system a hundred times each second over the entire age of the universe before you had a decent chance of ever finding it with $q_A < 10$. Even then, you would never find it with $q_A = 0$.

---

*There can be whole classes of states that are not accessible at all, perhaps because they have the wrong total energy. There can also be classes of states that are accessible only over time scales that are much longer than we are willing to wait. The concept of "accessible," like that of "macrostate," depends on the time scale under consideration. In the case of the Einstein solids, I'm assuming that all microstates with a given energy are accessible.

q_A	Ω_A	q_B	Ω_B	Ω_{total}
0	1	100	2.8×10^{81}	2.8×10^{81}
1	300	99	9.3×10^{80}	2.8×10^{83}
2	45150	98	3.1×10^{80}	1.4×10^{85}
3	4545100	97	1.0×10^{80}	4.6×10^{86}
4	3.4×10^{8}	96	3.3×10^{79}	1.1×10^{88}
⋮	⋮	⋮	⋮	⋮
59	2.2×10^{68}	41	3.1×10^{46}	6.8×10^{114}
60	1.3×10^{69}	40	5.3×10^{45}	6.9×10^{114}
61	7.7×10^{69}	39	8.8×10^{44}	6.8×10^{114}
⋮	⋮	⋮	⋮	⋮
100	1.7×10^{96}	0	1	1.7×10^{96}
				9.3×10^{115}

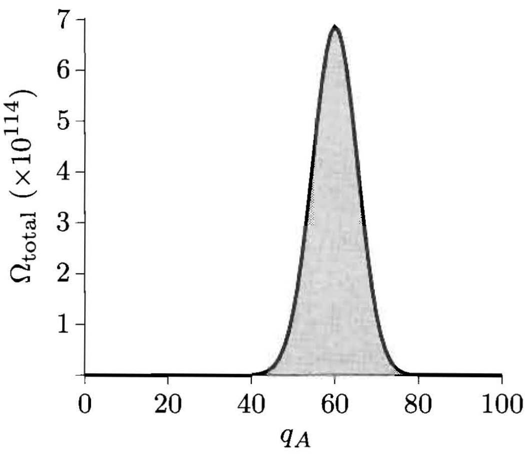


Figure 2.5. Macrostates and multiplicities of a system of two Einstein solids, with 300 and 200 oscillators respectively, sharing a total of 100 units of energy.

Suppose, however, that this system is *initially* in a state with q_A much less than 60; perhaps all the energy starts out in solid B. If you now wait a while for the energy to rearrange itself, then check again, you are more or less *certain* to find that energy has flowed from B to A. This system exhibits *irreversible* behavior: Energy flows spontaneously from B to A, but never (aside from small fluctuations around $q_A = 60$) from A to B. Apparently, we have discovered the physical explanation of *heat*: It is a *probabilistic* phenomenon, not absolutely certain, but extremely likely.

We have also stumbled upon a new law of physics: The spontaneous flow of energy *stops* when a system is at, or very near, its *most likely macrostate*, that is, the macrostate with the greatest multiplicity. This "law of increase of multiplicity" is one version of the famous **second law of thermodynamics**. Notice, though, that it's not a *fundamental* law at all—it's just a very strong statement about probabilities.

To make the statement stronger, and to be more realistic in general, we really should consider systems with not just a few hundred particles, but more like 10^{23}. Unfortunately, even a computer cannot calculate the number of ways of arranging 10^{23} units of energy among 10^{23} oscillators. Fortunately, there are some nice approximations we can make, to tackle this problem analytically. That is the subject of the following section.

Problem 2.8. Consider a system of two Einstein solids, A and B, each containing 10 oscillators, sharing a total of 20 units of energy. Assume that the solids are weakly coupled, and that the total energy is fixed.

- **(a)** How many different *macro*states are available to this system?
- **(b)** How many different *micro*states are available to this system?
- **(c)** Assuming that this system is in thermal equilibrium, what is the probability of finding all the energy in solid A?
- **(d)** What is the probability of finding exactly half of the energy in solid A?
- **(e)** Under what circumstances would this system exhibit irreversible behavior?

Problem 2.9. Use a computer to reproduce the table and graph in Figure 2.4: two Einstein solids, each containing three harmonic oscillators, with a total of six units of energy. Then modify the table and graph to show the case where one Einstein solid contains six harmonic oscillators and the other contains four harmonic oscillators (with the total number of energy units still equal to six). Assuming that all microstates are equally likely, what is the most probable macrostate, and what is its probability? What is the least probable macrostate, and what is its probability?

Problem 2.10. Use a computer to produce a table and graph, like those in this section, for the case where one Einstein solid contains 200 oscillators, the other contains 100 oscillators, and there are 100 units of energy in total. What is the most probable macrostate, and what is its probability? What is the least probable macrostate, and what is its probability?

Problem 2.11. Use a computer to produce a table and graph, like those in this section, for two interacting two-state paramagnets, each containing 100 elementary magnetic dipoles. Take a "unit" of energy to be the amount needed to flip a single dipole from the "up" state (parallel to the external field) to the "down" state (antiparallel). Suppose that the total number of units of energy, relative to the state with all dipoles pointing up, is 80; this energy can be shared in any way between the two paramagnets. What is the most probable macrostate, and what is its probability? What is the least probable macrostate, and what is its probability?

2.4 Large Systems

In the previous section we saw that, for a system of two interacting Einstein solids, each with a hundred or so oscillators, certain macrostates are *much* more probable than others. However, a significant fraction of the macrostates, roughly 20%, were still fairly probable. Next we'll look at what happens when the system is much larger, so that each solid contains, say, 10^{20} or more oscillators. My goal, by the end of this section, is to show you that out of all the macrostates, only a tiny fraction are reasonably probable. In other words, the multiplicity function becomes very sharp (see Figure 2.6). To analyze such large systems, however, we must first make a detour into the mathematics of very large numbers.

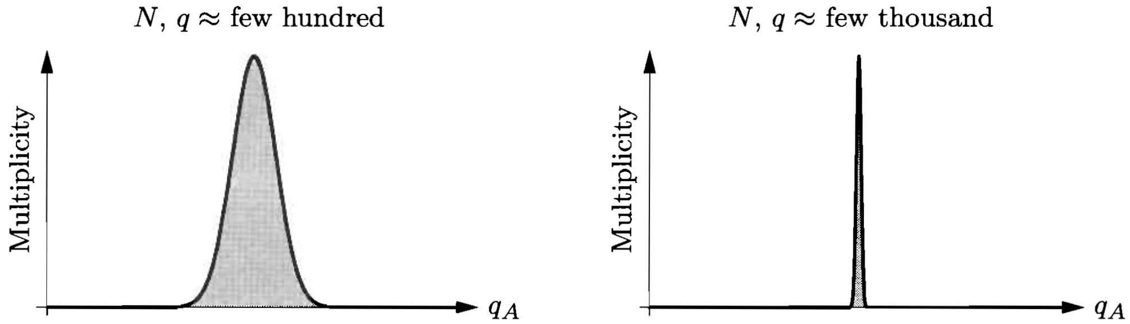


Figure 2.6. Typical multiplicity graphs for two interacting Einstein solids, containing a few hundred oscillators and energy units (left) and a few thousand (right). As the size of the system increases, the peak becomes very narrow relative to the full horizontal scale. For $N \approx q \approx 10^{20}$, the peak is much too sharp to draw.

Very Large Numbers

There are three kinds of numbers that commonly occur in statistical mechanics: small numbers, large numbers, and very large numbers.

Small numbers are small numbers, like 6, 23, and 42. You already know how to manipulate small numbers.

Large numbers are much larger than small numbers, and are frequently made by exponentiating small numbers. The most important large number in statistical mechanics is Avogadro's number, which is of order 10^{23}. The most important property of large numbers is that you can *add* a small number to a large number without changing it. For example,

$$10^{23} + 23 = 10^{23}. \tag{2.12}$$

(The only exception to this rule is when you plan to eventually subtract off the same large number: $10^{23} + 42 - 10^{23} = 42$.)

Very large numbers are even larger than large numbers, and can be made by exponentiating large numbers. An example would be* $10^{10^{23}}$. Very large numbers have the amazing property that you can *multiply* them by large numbers without changing them. For instance,

$$10^{10^{23}} \times 10^{23} = 10^{(10^{23}+23)} = 10^{10^{23}}, \tag{2.13}$$

by virtue of equation 2.12. This property takes some getting used to, but can be extremely convenient when manipulating very large numbers. (Again, there is an exception: When you plan to eventually *divide* by the same very large number, you need to keep track of any leftover factors.)

One common trick for manipulating very large numbers is to take the logarithm. This operation turns a *very* large number into an ordinary *large* number, which is much more familiar and can be manipulated more straightforwardly. Then at the end you can exponentiate to get back the very large number. I'll use this trick later in this section.

Problem 2.12. The natural logarithm function, ln, is defined so that $e^{\ln x} = x$ for any positive number x.

(a) Sketch a graph of the natural logarithm function.

(b) Prove the identities

$$\ln ab = \ln a + \ln b \qquad \text{and} \qquad \ln a^b = b \ln a.$$

(c) Prove that $\dfrac{d}{dx} \ln x = \dfrac{1}{x}$.

(d) Derive the useful approximation

$$\ln(1 + x) \approx x,$$

which is valid when $|x| \ll 1$. Use a calculator to check the accuracy of this approximation for $x = 0.1$ and $x = 0.01$.

---

*Note that x^{y^z} means $x^{(y^z)}$, not $(x^y)^z$.

Problem 2.13. Fun with logarithms.

(a) Simplify the expression $e^{a \ln b}$. (That is, write it in a way that doesn't involve logarithms.)

(b) Assuming that $b \ll a$, prove that $\ln(a+b) \approx (\ln a) + (b/a)$. (Hint: Factor out the a from the argument of the logarithm, so that you can apply the approximation of part (d) of the previous problem.)

Problem 2.14. Write $e^{10^{23}}$ in the form 10^x, for some x.

Stirling's Approximation

Our formulas for multiplicities involve "combinations," which involve factorials. To apply these formulas to *large* systems, we need a trick for evaluating factorials of large numbers. The trick is called **Stirling's approximation**:

$$N! \approx N^N e^{-N} \sqrt{2\pi N}. \tag{2.14}$$

This approximation is accurate in the limit where $N \gg 1$. Let me try to explain why.

The quantity $N!$ is the product of N factors, from 1 up to N. A very crude approximation would be to replace each of the N factors in the factorial by N, so $N! \approx N^N$. This is a gross *over*estimate, since nearly all of the N factors in $N!$ are actually smaller than N. It turns out that, on average, each factor is effectively smaller by a factor of e:

$$N! \approx \left(\frac{N}{e}\right)^N = N^N e^{-N}. \tag{2.15}$$

This is still off by a large factor, roughly $\sqrt{2\pi N}$. But if N is a large number, then $N!$ is a *very* large number, and often this correction factor (which is only a *large* number) can be omitted.

If all you care about is the *logarithm* of $N!$, then equation 2.15 is usually good enough. Another way to write it is

$$\ln N! \approx N \ln N - N. \tag{2.16}$$

It's fun to test Stirling's approximation on some not-very-large numbers, using a calculator or a computer. Table 2.3 shows a sampling of results. As you can see, N does not have to be particularly large before Stirling's approximation becomes useful. Equation 2.14 is quite accurate even for $N = 10$, while equation 2.16 is quite accurate for $N = 100$ (if all you care about is the logarithm).

For a *derivation* of Stirling's approximation, see Appendix B.

N	$N!$	$N^N e^{-N}\sqrt{2\pi N}$	Error	$\ln N!$	$N\ln N - N$	Error
1	1	.922	7.7%	0	-1	∞
10	3628800	3598696	.83%	15.1	13.0	13.8%
100	9×10^{157}	9×10^{157}	.083%	364	360	.89%

Table 2.3. Comparison of Stirling's approximation (equations 2.14 and 2.16) to exact values for $N = 1$, 10, and 100.

Problem 2.15. Use a pocket calculator to check the accuracy of Stirling's approximation for $N = 50$. Also check the accuracy of equation 2.16 for $\ln N!$.

Problem 2.16. Suppose you flip 1000 coins.

- (a) What is the probability of getting *exactly* 500 heads and 500 tails? (Hint: First write down a formula for the total number of possible outcomes. Then, to determine the "multiplicity" of the 500-500 "macrostate," use Stirling's approximation. If you have a fancy calculator that makes Stirling's approximation unnecessary, multiply all the numbers in this problem by 10, or 100, or 1000, until Stirling's approximation becomes necessary.)
- (b) What is the probability of getting exactly 600 heads and 400 tails?

Multiplicity of a Large Einstein Solid

Armed with Stirling's approximation, let me now estimate the multiplicity of an Einstein solid containing a *large* number of oscillators and energy units. Rather than working it out in complete generality, I'll consider only the case $q \gg N$, when there are many more energy units than oscillators. (This is the "high-temperature" limit.)

I'll start with the exact formula:

$$\Omega(N,q) = \binom{q+N-1}{q} = \frac{(q+N-1)!}{q!\,(N-1)!} \approx \frac{(q+N)!}{q!\,N!}. \tag{2.17}$$

I'm making the last approximation because the difference between $N!$ and $(N-1)!$ is only a large factor (N), which is insignificant in a *very* large number like Ω. Next I'll take the natural logarithm and apply Stirling's approximation in the form 2.16:

$$\begin{aligned}\ln\Omega &= \ln\left(\frac{(q+N)!}{q!\,N!}\right)\\ &= \ln(q+N)! - \ln q! - \ln N!\\ &\approx (q+N)\ln(q+N) - (q+N) - q\ln q + q - N\ln N + N\\ &= (q+N)\ln(q+N) - q\ln q - N\ln N.\end{aligned} \tag{2.18}$$

So far I haven't assumed that $q \gg N$—only that both q and N are large. But now let me manipulate the first logarithm as in Problem 2.13:

$$\begin{aligned}\ln(q+N) &= \ln\left[q\left(1+\frac{N}{q}\right)\right]\\ &= \ln q + \ln\left(1+\frac{N}{q}\right)\\ &\approx \ln q + \frac{N}{q}.\end{aligned} \tag{2.19}$$

The last step follows from the Taylor expansion of the logarithm, $\ln(1+x) \approx x$ for $|x| \ll 1$. Plugging this result into equation 2.18 and canceling the $q\ln q$ terms, we obtain

$$\ln\Omega \approx N\ln\frac{q}{N} + N + \frac{N^2}{q}. \tag{2.20}$$

The last term becomes negligible compared to the others in the limit $q \gg N$. Exponentiating the first two terms gives

$$\Omega(N,q) \approx e^{N\ln(q/N)}e^N = \left(\frac{eq}{N}\right)^N \qquad \text{(when } q \gg N). \tag{2.21}$$

This formula is nice and simple, but it's bizarre. The exponent is a *large* number, so Ω is a *very* large number, as we already knew. Furthermore, if you increase either N or q by just a little bit, Ω will increase by a *lot*, due to the large exponent N.

Problem 2.17. Use the methods of this section to derive a formula, similar to equation 2.21, for the multiplicity of an Einstein solid in the "low-temperature" limit, $q \ll N$.

Problem 2.18. Use Stirling's approximation to show that the multiplicity of an Einstein solid, for any large values of N and q, is approximately

$$\Omega(N,q) \approx \frac{\left(\frac{q+N}{q}\right)^q \left(\frac{q+N}{N}\right)^N}{\sqrt{2\pi q(q+N)/N}}.$$

The square root in the denominator is merely large, and can often be neglected. However, it is needed in Problem 2.22. (Hint: First show that $\Omega = \frac{N}{q+N}\frac{(q+N)!}{q!\,N!}$. Do not neglect the $\sqrt{2\pi N}$ in Stirling's approximation.)

Problem 2.19. Use Stirling's approximation to find an approximate formula for the multiplicity of a two-state paramagnet. Simplify this formula in the limit $N_\downarrow \ll N$ to obtain $\Omega \approx (Ne/N_\downarrow)^{N_\downarrow}$. This result should look very similar to your answer to Problem 2.17; explain why these two systems, in the limits considered, are essentially the same.

Sharpness of the Multiplicity Function

Finally we're ready to return to the issue raised at the beginning of this section: For a system of *two* large, interacting Einstein solids, just how skinny is the peak in the multiplicity function?

For simplicity, let me assume that each solid has N oscillators. I'll call the total number of energy units simply q (instead of q_{total}) for brevity, and I'll assume that this is much larger than N, so we can use formula 2.21. Then the multiplicity of the combined system, for any given macrostate, is

$$\Omega = \left(\frac{eq_A}{N}\right)^N \left(\frac{eq_B}{N}\right)^N = \left(\frac{e}{N}\right)^{2N} (q_A q_B)^N, \tag{2.22}$$

where q_A and q_B are the numbers of energy units in solids A and B. (Note that $q_A + q_B$ must equal q.)

If you graph equation 2.22 as a function of q_A, it will have a very sharp peak at $q_A = q/2$, where the energy is distributed equally between the solids. The height of this peak is a very large number:

$$\Omega_{\text{max}} = \left(\frac{e}{N}\right)^{2N} \left(\frac{q}{2}\right)^{2N}. \tag{2.23}$$

I want to know what the graph looks like *near* this peak, so let me set

$$q_A = \frac{q}{2} + x, \qquad q_B = \frac{q}{2} - x, \tag{2.24}$$

where x can be any number that is much smaller than q (but still possibly quite large). Plugging these expressions into equation 2.22 gives

$$\Omega = \left(\frac{e}{N}\right)^{2N} \left[\left(\frac{q}{2}\right)^2 - x^2\right]^N. \tag{2.25}$$

To simplify the second factor, I'll take its logarithm and manipulate it as I did in equation 2.19:

$$\begin{aligned}
\ln\left[\left(\frac{q}{2}\right)^2 - x^2\right]^N &= N \ln\left[\left(\frac{q}{2}\right)^2 - x^2\right] \\
&= N \ln\left[\left(\frac{q}{2}\right)^2 \left(1 - \left(\frac{2x}{q}\right)^2\right)\right] \\
&= N\left[\ln\left(\frac{q}{2}\right)^2 + \ln\left(1 - \left(\frac{2x}{q}\right)^2\right)\right] \\
&\approx N\left[\ln\left(\frac{q}{2}\right)^2 - \left(\frac{2x}{q}\right)^2\right].
\end{aligned} \tag{2.26}$$

Now I can exponentiate the last expression and plug this back into equation 2.25:

$$\Omega = \left(\frac{e}{N}\right)^{2N} e^{N\ln(q/2)^2} e^{-N(2x/q)^2} = \Omega_{\text{max}} \cdot e^{-N(2x/q)^2}. \tag{2.27}$$

A function of this form is called a **Gaussian**; it has a peak at $x = 0$ and a sharp fall-off on either side, as shown in Figure 2.7. The multiplicity falls off to $1/e$ of its maximum value when

$$N\left(\frac{2x}{q}\right)^2 = 1 \qquad \text{or} \qquad x = \frac{q}{2\sqrt{N}}. \tag{2.28}$$

This is actually a rather large number. But if $N = 10^{20}$, it's only one part in ten billion of the entire scale of the graph! On the scale used in the figure, where the

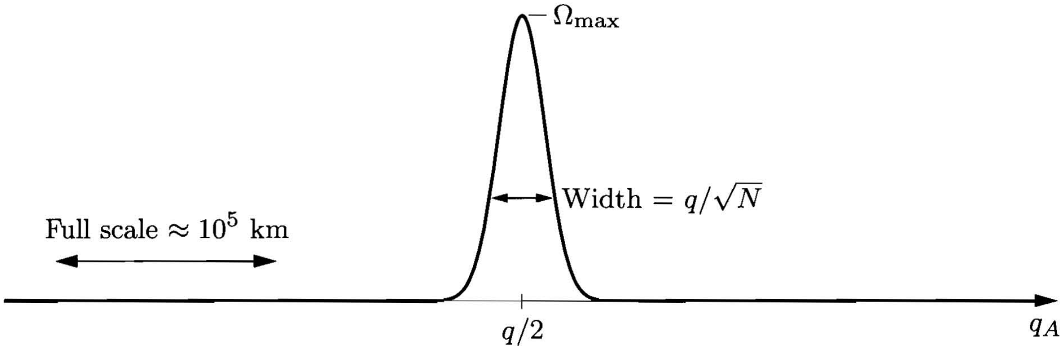


Figure 2.7. Multiplicity of a system of two large Einstein solids with many energy units per oscillator (high-temperature limit). Only a tiny fraction of the full horizontal scale is shown.

width of the peak is about 1 cm, the full scale of the graph would have to stretch 10^{10} cm, or 100,000 km—more than twice around the earth. And near the edge of the page, where x is only ten times larger than $q/2\sqrt{N}$, the multiplicity is less than its maximum value by a factor of $e^{-100} \approx 10^{-44}$.

This result tells us that, when two large Einstein solids are in thermal equilibrium with each other, any random fluctuations away from the most likely macrostate will be *utterly unmeasurable.* To measure such fluctuations we would have to measure the energy to an accuracy of ten significant figures. Once the system has had time to come to thermal equilibrium, so that all *micro*states are equally probable, we might as well assume that it is in its *most* likely macrostate. The limit where a system becomes infinitely large, so that fluctuations away from the most likely macrostate never occur, is called the **thermodynamic limit**.

Problem 2.20. Suppose you were to shrink Figure 2.7 until the entire horizontal scale fits on the page. How wide would the peak be?

Problem 2.21. Use a computer to plot formula 2.22 directly, as follows. Define $z = q_A/q$, so that $(1-z) = q_B/q$. Then, aside from an overall constant that we'll ignore, the multiplicity function is $[4z(1-z)]^N$, where z ranges from 0 to 1 and the factor of 4 ensures that the height of the peak is equal to 1 for any N. Plot this function for $N = 1$, 10, 100, 1000, and 10,000. Observe how the width of the peak decreases as N increases.

Problem 2.22. This problem gives an alternative approach to estimating the width of the peak of the multiplicity function for a system of two large Einstein solids.

(a) Consider two identical Einstein solids, each with N oscillators, in thermal contact with each other. Suppose that the total number of energy units in the combined system is exactly $2N$. How many different macrostates (that is, possible values for the total energy in the first solid) are there for this combined system?

(b) Use the result of Problem 2.18 to find an approximate expression for the total number of microstates for the combined system. (Hint: Treat the combined system as a single Einstein solid. Do *not* throw away factors of "large" numbers, since you will eventually be dividing two "very large" numbers that are nearly equal. *Answer*: $2^{4N}/\sqrt{8\pi N}$.)

(c) The most likely macrostate for this system is (of course) the one in which the energy is shared equally between the two solids. Use the result of Problem 2.18 to find an approximate expression for the multiplicity of this macrostate. (*Answer*: $2^{4N}/(4\pi N)$.)

(d) You can get a rough idea of the "sharpness" of the multiplicity function by comparing your answers to parts (b) and (c). Part (c) tells you the height of the peak, while part (b) tells you the total area under the entire graph. As a very crude approximation, pretend that the peak's shape is rectangular. In this case, how wide would it be? Out of all the macrostates, what fraction have reasonably large probabilities? Evaluate this fraction numerically for the case $N = 10^{23}$.

Problem 2.23. Consider a two-state paramagnet with 10^{23} elementary dipoles, with the total energy fixed at zero so that exactly half the dipoles point up and half point down.

- **(a)** How many microstates are "accessible" to this system?
- **(b)** Suppose that the microstate of this system changes a billion times per second. How many microstates will it explore in ten billion years (the age of the universe)?
- **(c)** Is it correct to say that, if you wait long enough, a system will eventually be found in every "accessible" microstate? Explain your answer, and discuss the meaning of the word "accessible."

Problem 2.24. For a single *large* two-state paramagnet, the multiplicity function is very sharply peaked about $N_\uparrow = N/2$.

- **(a)** Use Stirling's approximation to estimate the height of the peak in the multiplicity function.
- **(b)** Use the methods of this section to derive a formula for the multiplicity function in the vicinity of the peak, in terms of $x \equiv N_\uparrow - (N/2)$. Check that your formula agrees with your answer to part (a) when $x = 0$.
- **(c)** How wide is the peak in the multiplicity function?
- **(d)** Suppose you flip 1,000,000 coins. Would you be surprised to obtain 501,000 heads and 499,000 tails? Would you be surprised to obtain 510,000 heads and 490,000 tails? Explain.

Problem 2.25. The mathematics of the previous problem can also be applied to a one-dimensional **random walk**: a journey consisting of N steps, all the same size, each chosen randomly to be either forward or backward. (The usual mental image is that of a drunk stumbling along an alley.)

- **(a)** Where are you *most* likely to find yourself, after the end of a long random walk?
- **(b)** Suppose you take a random walk of 10,000 steps (say each a yard long). About how far from your starting point would you expect to be at the end?
- **(c)** A good example of a random walk in nature is the **diffusion** of a molecule through a gas; the average step length is then the mean free path, as computed in Section 1.7. Using this model, and neglecting any small numerical factors that might arise from the varying step size and the multidimensional nature of the path, estimate the expected net displacement of an air molecule (or perhaps a carbon monoxide molecule traveling through air) in one second, at room temperature and atmospheric pressure. Discuss how your estimate would differ if the elapsed time or the temperature were different. Check that your estimate is consistent with the treatment of diffusion in Section 1.7.

It all works because Avogadro's number is closer to infinity than to 10.

—Ralph Baierlein, *American Journal of Physics* **46**, 1045 (1978). Copyright 1978, American Association of Physics Teachers. Reprinted with permission.

2.5 The Ideal Gas

The conclusion of the previous section—that only a *tiny* fraction of the macrostates of a large interacting system have reasonably large probabilities—applies to many other systems besides Einstein solids. In fact, it is true for essentially *any* pair of interacting objects, provided that the number of particles and the number of energy units are both "large." In this section I'll argue that it is true for ideal gases.

An ideal gas is more complicated than an Einstein solid, because its multiplicity depends on its volume as well as its total energy and number of particles. Furthermore, when two gases interact, they can often expand and contract, and even exchange molecules, in addition to exchanging energy. We will still find, however, that the multiplicity function for two interacting gases is very sharply peaked around a relatively small subset of macrostates.

Multiplicity of a Monatomic Ideal Gas

For simplicity, I'll consider only a monatomic ideal gas, like helium or argon. I'll begin with a gas consisting of just one molecule, then work up to the general case of N molecules.

So suppose we have a single gas atom, with total energy U, in a container of volume V. What is the multiplicity of this system? That is, how many microstates could the molecule be in, given the fixed values of U and V?

Well, a container with twice the volume offers twice as many states to a molecule, so the multiplicity should be proportional to V. Also, the more different momentum vectors the molecule can have, the more states are available, so the multiplicity should also be proportional to the "volume" of available **momentum space**. (Momentum space is an imaginary "space" in which the axes are p_x, p_y, and p_z. Each "point" in momentum space corresponds to a momentum vector for the particle.) So let me write schematically

$$\Omega_1 \propto V \cdot V_p, \tag{2.29}$$

where V is the volume of ordinary space (or **position space**), V_p is the volume of momentum space, and the 1 subscript indicates that this is for a gas of just one molecule.

This formula for Ω_1 is still pretty ambiguous. One problem is in determining the available volume of momentum space, V_p. Since the molecule's kinetic energy must equal U, there is a constraint:

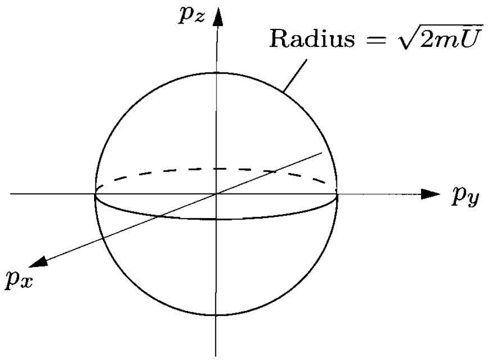
$$U = \frac{1}{2}m\left(v_x^2 + v_y^2 + v_z^2\right) = \frac{1}{2m}\left(p_x^2 + p_y^2 + p_z^2\right). \tag{2.30}$$

This equation can also be written

$$p_x^2 + p_y^2 + p_z^2 = 2mU, \tag{2.31}$$

which defines the surface of a *sphere* in momentum space with radius $\sqrt{2mU}$ (see Figure 2.8). The "volume" of momentum space is really the *surface area* of this sphere (perhaps multiplied by a small thickness if U is allowed to fluctuate somewhat).

Figure 2.8. A sphere in momentum space with radius $\sqrt{2mU}$. If a molecule has energy U, its momentum vector must lie somewhere on the surface of this sphere.



The other problem with equation 2.29 is in determining the constant of proportionality. While it seems pretty clear that Ω_1 must be *proportional* to the volumes of position space and momentum space, how can we possibly *count* the various microstates to get a finite number for the multiplicity? It would seem that the number of allowed microstates, even for a gas of just one molecule, is infinite.

To actually count the number of microstates we must invoke quantum mechanics. (For a systematic overview of quantum mechanics, see Appendix A.) In quantum mechanics, the state of a system is described by a *wavefunction*, which is spread out in both position space and momentum space. The less spread out the wavefunction is in position space, the more spread out it must be in momentum space, and vice versa. This is the famous **Heisenberg uncertainty principle**:

$$(\Delta x)(\Delta p_x) \approx h, \tag{2.32}$$

where Δx is the spread in x, Δp_x is the spread in p_x, and h is Planck's constant. (The product of Δx and Δp_x can also be *more* than h, but we are interested in wavefunctions that specify the position and momentum as precisely as possible.) The same limitation applies to y and p_y, and to z and p_z.

Even in quantum mechanics, the number of allowed wavefunctions is infinite. But the number of *independent* wavefunctions (in a technical sense that's defined in Appendix A) is finite, if the total available position space and momentum space are limited. I like to picture it as in Figure 2.9. In this one-dimensional example, the number of distinct position states is $L/(\Delta x)$, while the number of distinct

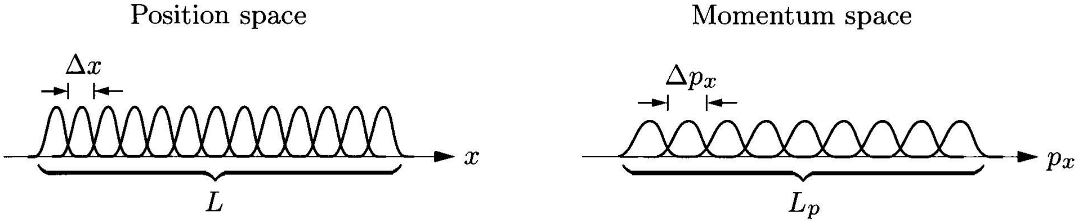


Figure 2.9. A number of "independent" position states and momentum states for a quantum-mechanical particle moving in one dimension. If we make the wavefunctions narrower in position space, they become wider in momentum space, and vice versa.

momentum states is $L_p/(\Delta p_x)$. The total number of distinct states is the product,

$$\frac{L}{\Delta x}\frac{L_p}{\Delta p_x} = \frac{L\,L_p}{h}, \tag{2.33}$$

according to the uncertainty principle. In three dimensions, the lengths become volumes and there are three factors of h:

$$\Omega_1 = \frac{V\,V_p}{h^3}. \tag{2.34}$$

This "derivation" of the constant of proportionality in Ω_1 is admittedly not very rigorous. I certainly haven't proved that there are no further factors of 2 or π in equation 2.34. If you prefer, just think of the result in terms of dimensional analysis: The multiplicity must be a unitless number, and you can easily show that h^3 has just the right units to cancel the units of V and V_p.*

So much for a gas of one molecule. If we add a second molecule, we need a factor of the form of equation 2.34 for each molecule, and we multiply them together because for *each* state of molecule 1, there are Ω_1 states for molecule 2. Well, not quite. The V_p factors are more complicated, since only the *total* energy of the two molecules is constrained. Equation 2.31 now becomes

$$p_{1x}^2 + p_{1y}^2 + p_{1z}^2 + p_{2x}^2 + p_{2y}^2 + p_{2z}^2 = 2mU, \tag{2.35}$$

assuming that both molecules have the same mass. This equation defines the surface of a six-dimensional "hypersphere" in six-dimensional momentum space. I can't visualize it, but one can still compute its "surface area" and call that the total volume of allowed momentum space for the two molecules.

So the multiplicity function for an ideal gas of two molecules should be

$$\Omega_2 = \frac{V^2}{h^6} \times (\text{area of momentum hypersphere}). \tag{2.36}$$

This formula *is* correct, but only if the two molecules are *distinguishable* from each other. If they're *indistinguishable*, then we've overcounted the microstates by a factor of 2, since interchanging the molecules with each other does not give us a distinct state (see Figure 2.10).† Thus the multiplicity for a gas of two *indistinguishable* molecules is

$$\Omega_2 = \frac{1}{2}\frac{V^2}{h^6} \times (\text{area of momentum hypersphere}). \tag{2.37}$$

---

*Don't worry about the fact that V_p is really a surface area, not a volume. We can always allow the sphere in momentum space to have a tiny thickness, and multiply its area by this thickness to get something with units of momentum cubed. When we get to a gas of N molecules, the multiplicity will be such a huge number that it doesn't matter if we're off a little in the units.

†This argument assumes that the individual states of the two molecules are always different. The two molecules *could* be in a state where they both have the same position and the same momentum, and such a state is not double-counted in equation 2.36. Unless the gas is *very* dense, however, such states hardly ever occur.

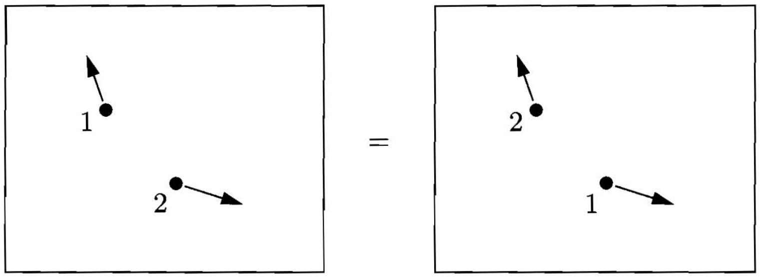


Figure 2.10. In a gas of two identical molecules, interchanging the states of the molecules leaves the system in the same state as before.

For an ideal gas of N indistinguishable molecules, the multiplicity function contains N factors of V, divided by $3N$ factors of h. The factor that compensates for the overcounting is $1/N!$, the number of ways of interchanging the molecules. And the momentum-space factor is the "surface area" of a $3N$-dimensional hypersphere whose radius is (still) $\sqrt{2mU}$:

$$\Omega_N = \frac{1}{N!}\frac{V^N}{h^{3N}} \times (\text{area of momentum hypersphere}). \tag{2.38}$$

To make this result more explicit, we need a general formula for the "surface area" of a d-dimensional hypersphere of radius r. For $d = 2$, the "area" is just the circumference of a circle, $2\pi r$. For $d = 3$, the answer is $4\pi r^2$. For general d, the answer should be proportional to r^{d-1}, but the coefficient is not easy to guess. The full formula is

$$\text{“area”} = \frac{2\pi^{d/2}}{(\frac{d}{2}-1)!} r^{d-1}. \tag{2.39}$$

This formula is derived in Appendix B. For now, you can at least check the coefficient for the case $d = 2$. To check it for $d = 3$, you need to know that $(1/2)! = \sqrt{\pi}/2$.

Plugging equation 2.39 (with $d = 3N$ and $r = \sqrt{2mU}$) into equation 2.38, we obtain

$$\Omega_N = \frac{1}{N!}\frac{V^N}{h^{3N}}\frac{2\pi^{3N/2}}{(\frac{3N}{2}-1)!}(\sqrt{2mU})^{3N-1} \approx \frac{1}{N!}\frac{V^N}{h^{3N}}\frac{\pi^{3N/2}}{(3N/2)!}(\sqrt{2mU})^{3N}. \tag{2.40}$$

In the last expression I've thrown away some large factors, which is ok since Ω_N is a very large number.*

This formula for the multiplicity of a monatomic ideal gas is a mess, but its dependence on U and V is pretty simple:

$$\Omega(U, V, N) = f(N)V^N U^{3N/2}, \tag{2.41}$$

where $f(N)$ is a complicated function of N.

---

*If you're not happy with my sloppy derivation of equation 2.40, please be patient. In Section 6.7 I'll do a much better job, using a very different method.

Notice that the exponent on U in formula 2.41 is 1/2 times the total number of degrees of freedom ($3N$) in the monatomic gas. The same is true of the multiplicity of an Einstein solid in the high-temperature limit, equation 2.21. These results are special cases of a more general theorem: For any system with only quadratic "degrees of freedom," having so many units of energy that energy quantization is unnoticeable, the multiplicity is proportional to $U^{Nf/2}$, where Nf is the total number of degrees of freedom. A general proof of this theorem is given in Stowe (1984).

Problem 2.26. Consider an ideal monatomic gas that lives in a two-dimensional universe ("flatland"), occupying an area A instead of a volume V. By following the same logic as above, find a formula for the multiplicity of this gas, analogous to equation 2.40.

Interacting Ideal Gases

Suppose now that we have *two* ideal gases, separated by a partition that allows energy to pass through (see Figure 2.11). If each gas has N molecules (of the same species), then the total multiplicity of this system is

$$\Omega_{\text{total}} = [f(N)]^2 (V_A V_B)^N (U_A U_B)^{3N/2}. \tag{2.42}$$

This expression has essentially the same form as the corresponding result for a pair of Einstein solids (equation 2.22): Both energies are raised to a *large* exponent. Following exactly the same reasoning as in Section 2.4, we can conclude that the multiplicity function, plotted as a function of U_A, has a *very* sharp peak:

$$\text{width of peak} = \frac{U_{\text{total}}}{\sqrt{3N/2}}. \tag{2.43}$$

Provided that N is large, only a *tiny* fraction of the macrostates have a reasonable chance of occurring, assuming that the system is in equilibrium.

In addition to exchanging energy, we could allow the gases to exchange *volume*; that is, we could allow the partition to move back and forth, as one gas expands and the other is compressed. In this case we can apply exactly the same argument to volume that we just applied to energy. The multiplicity, plotted as a function

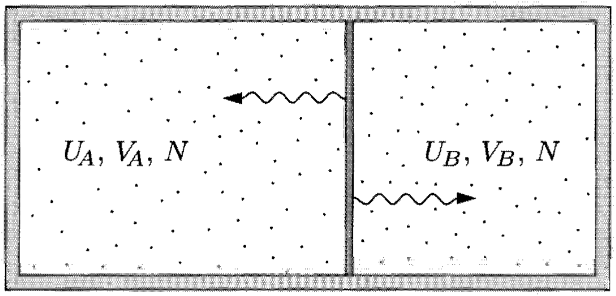


Figure 2.11. Two ideal gases, each confined to a fixed volume, separated by a partition that allows energy to pass through. The total energy of the two gases is fixed.

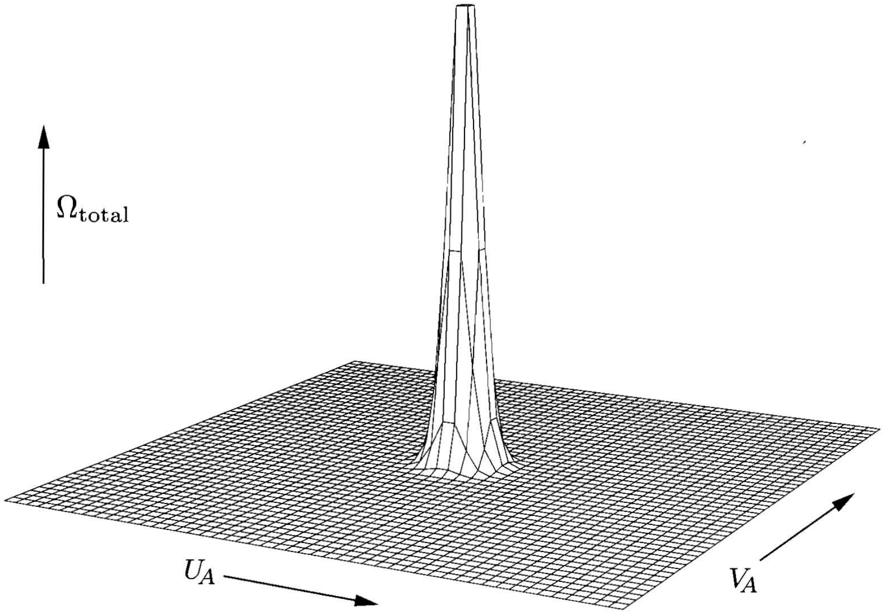


Figure 2.12. Multiplicity of a system of two ideal gases, as a function of the energy and volume of gas A (with the total energy and total volume held fixed). If the number of molecules in each gas is large, the full horizontal scale would stretch far beyond the edge of the page.

of V_A, again has a very sharp peak:

$$\text{width of peak} = \frac{V_{\text{total}}}{\sqrt{N}}. \tag{2.44}$$

So again, the equilibrium macrostate is essentially determined, to within a tiny fraction of the total volume available (if N is large). In Figure 2.12 I've plotted Ω_{total} as a function of *both* U_A and V_A. Like Figure 2.7, this graph shows only a tiny fraction of the full range of U_A and V_A values. For $N = 10^{20}$, if the full scale were compressed to fit on this page, the spike would be narrower than an atom.

Instead of allowing the partition to move, we could just poke holes in it and let the *molecules* move back and forth between the two sides. Then, to find the equilibrium macrostate, we would want to look at the behavior of Ω_{total} as a function of N_A and U_A. From equation 2.40, you can see that the analysis would be more difficult in this case. But once again, we would find a very sharp peak in the graph, indicating that the equilibrium macrostate is fixed to a very high precision. (As you might expect, the equilibrium macrostate is the one for which the density is the same on both sides of the partition.)

Sometimes you can calculate probabilities of various arrangements of molecules just by looking at the volume dependence of the multiplicity function (2.41). For instance, suppose we want to know the probability of finding the configuration shown in Figure 2.13, where *all* the molecules in a container of gas are somewhere in the left half. This arrangement is just a macrostate with the same energy and number of molecules, but half the original volume. Looking at equation 2.41, we see that replacing V by $V/2$ reduces the multiplicity by a factor of 2^N. In other

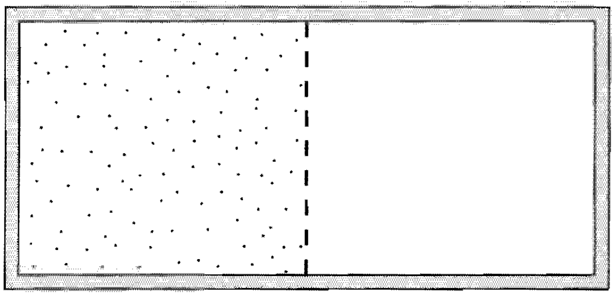


Figure 2.13. A very unlikely arrangement of gas molecules.

words, out of all the allowed microstates, only one in 2^N has all the molecules in the left half. Thus, the probability of this arrangement is 2^{-N}. Even for $N = 100$, this is less than 10^{-30}, so you would have to check a trillion times per second for the age of the universe before finding such an arrangement even once. For $N = 10^{23}$, the probability is a *very* small number.

Problem 2.27. Rather than insisting that all the molecules be in the left half of a container, suppose we only require that they be in the leftmost 99% (leaving the remaining 1% completely empty). What is the probability of finding such an arrangement if there are 100 molecules in the container? What if there are 10,000 molecules? What if there are 10^{23}?

2.6 Entropy

We have now seen that, for a variety of systems, particles and energy tend to rearrange themselves until the multiplicity is at (or very near) its maximum value. In fact, this conclusion seems to be true* for *any* system, provided that it contains enough particles and units of energy for the statistics of very large numbers to apply:

> Any large system in equilibrium will be found in the macrostate with the greatest multiplicity (aside from fluctuations that are normally too small to measure).

This is just a more general statement of the **second law of thermodynamics**. Another way to say it is simply:

> Multiplicity tends to increase.

Even though this law is not "fundamental" (since I essentially derived it by looking at probabilities), I'll treat it as fundamental from now on. If you just remember to look for the macrostate with greatest multiplicity, you can pretty much forget about calculating what the actual probabilities are.

---

*As far as I'm aware, nobody has ever *proved* that it is true for *all* large systems. Perhaps an exception lurks out there somewhere. But the experimental successes of thermodynamics indicate that exceptions must be exceedingly rare.

Since multiplicities tend to be *very* large numbers, which are very cumbersome to work with, we will find it convenient from now on to work with the natural logarithm of the multiplicity instead of the multiplicity itself. For historical reasons, we will also multiply by a factor of Boltzmann's constant. This gives us a quantity called the **entropy**, denoted S:

$$S \equiv k \ln \Omega. \tag{2.45}$$

In words, entropy is just the logarithm of the number of ways of arranging things in the system (times Boltzmann's constant). The logarithm turns a very large number, the multiplicity, into an ordinary large number. If you want to *understand* entropy, my advice is to ignore the factor of k and just think of entropy as a unitless quantity, $\ln \Omega$. When we include the factor of k, however, S has units of energy divided by temperature, or J/K in the SI system. I'll explain the usefulness of these units in Chapter 3.

As a first example, let's go back to the case of a large Einstein solid with N oscillators, q units of energy, and $q \gg N$. Since $\Omega = (eq/N)^N$,

$$S = k \ln(eq/N)^N = Nk[\ln(q/N) + 1]. \tag{2.46}$$

So if $N = 10^{22}$ and $q = 10^{24}$,

$$S = Nk \cdot (5.6) = (5.6 \times 10^{22})k = 0.77 \text{ J/K}. \tag{2.47}$$

Notice also that increasing either q or N increases the entropy of an Einstein solid (though not in direct proportion).

Generally, the more particles there are in a system, and the more energy it contains, the greater its multiplicity and its entropy. Besides adding particles and energy, you can increase the entropy of a system by letting it expand into a larger space, or breaking large molecules apart into small ones, or mixing together substances that were once separate. In each of these cases, the total number of possible arrangements increases.

Some people find it helpful to think of entropy intuitively as being roughly synonymous with "disorder." Whether this idea is accurate, however, depends on exactly what you consider to be disorderly. Most people would agree that a shuffled deck of cards is more disorderly than a sorted deck, and indeed, shuffling increases the entropy because it increases the number of possible arrangements.* However, many people would say that a glass of crushed ice appears more disorderly than a glass of an equal amount of water. In this case, though, the water has much more entropy, since there are so many more ways of arranging the molecules, and so many more ways of arranging the larger amount of energy among them.

---

*This example is actually somewhat controversial: Some physicists would not count these rearrangements into the thermodynamic entropy because cards don't ordinarily rearrange themselves without outside help. Personally, I see no point in being so picky. At worst, my somewhat broad definition of entropy is harmless, because the amount of entropy in dispute is negligible compared to other forms of entropy.

One nice property of entropy is that the total entropy of a composite system is the *sum* of the entropies of its parts. For instance, if there are two parts, A and B, then

$$S_{\text{total}} = k \ln \Omega_{\text{total}} = k \ln(\Omega_A \Omega_B) = k \ln \Omega_A + k \ln \Omega_B = S_A + S_B. \tag{2.48}$$

I'm assuming here that the macrostates of systems A and B have been specified separately. If these systems can interact, then those macrostates can fluctuate over time, and to compute the entropy over long time scales we should compute Ω_{total} by summing over *all* macrostates for the two systems. Entropy, like multiplicity, is a function of the number of *accessible* microstates, and this number depends on the time scale under consideration. However, in practice, this distinction rarely matters. If we just assume that the composite system is in its *most* likely macrostate, we get essentially the same entropy as if we sum over *all* macrostates (see Problems 2.29 and 2.30).

Since the natural logarithm is a monotonically increasing function of its argument, a macrostate with higher multiplicity also has higher entropy. Therefore we can restate the **second law of thermodynamics** as follows:

> Any large system in equilibrium will be found in the macrostate with the greatest entropy (aside from fluctuations that are normally too small to measure).

Or more briefly:

> Entropy tends to increase.

Note, however, that a graph of entropy vs. some variable (such as U_A or V_A) that is allowed to fluctuate will generally *not* have a sharp peak. Taking the logarithm smooths out the peak that was present in the multiplicity function. Of course this does not affect our conclusions in the least; it is still true that fluctuations away from the macrostate of greatest entropy will be negligible for any reasonably large system.

Although "spontaneous" processes always occur because of a net increase in entropy, you might wonder whether human intervention could bring about a net decrease in entropy. Common experience seems to suggest that the answer is yes: Anyone can easily turn all the coins in a collection heads-up, or sort a shuffled deck of cards, or clean up a messy room. However, the decreases in entropy in these situations are extremely tiny, while the entropy created by the metabolism of food in our bodies (as we take energy out of chemical bonds and dump most of it into the environment as thermal energy) is always substantial. As far as we can tell, our bodies are just as subject to the laws of thermodynamics as are inanimate objects. So no matter what you do to decrease the entropy in one place, you're bound to create at least as much entropy somewhere else.

Even if *we* can't decrease the total entropy of the universe, isn't it possible that someone (or something) else could? In 1867 James Clerk Maxwell posed this question, wondering whether a "very observant and neat-fingered being"* couldn't

---

*Quoted in Leff and Rex (1990), p. 5.

deflect fast-moving molecules in one direction and slow-moving molecules in another, thereby causing heat to flow from a cold object to a hot one. William Thomson later named this mythical creature *Maxwell's Demon*, and physicists and philosophers have been trying to exorcise it ever since. Countless designs for mechanical "demons" have been drafted, and all have been proven ineffective. Even a hypothetical "intelligent" demon, it turns out, must create entropy as it processes the information needed to sort molecules. Although thinking about demons has taught us much about entropy since Maxwell's time, the verdict seems to be that not even a demon can violate the second law of thermodynamics.

Problem 2.28. How many possible arrangements are there for a deck of 52 playing cards? (For simplicity, consider only the order of the cards, not whether they are turned upside-down, etc.) Suppose you start with a sorted deck and shuffle it repeatedly, so that all arrangements become "accessible." How much entropy do you create in the process? Express your answer both as a pure number (neglecting the factor of k) and in SI units. Is this entropy significant compared to the entropy associated with arranging thermal energy among the molecules in the cards?

Problem 2.29. Consider a system of two Einstein solids, with $N_A = 300$, $N_B = 200$, and $q_{\text{total}} = 100$ (as discussed in Section 2.3). Compute the entropy of the most likely macrostate and of the least likely macrostate. Also compute the entropy over long time scales, assuming that *all* microstates are accessible. (Neglect the factor of Boltzmann's constant in the definition of entropy; for systems this small it is best to think of entropy as a pure number.)

Problem 2.30. Consider again the system of two large, identical Einstein solids treated in Problem 2.22.

- **(a)** For the case $N = 10^{23}$, compute the entropy of this system (in terms of Boltzmann's constant), assuming that *all* of the microstates are allowed. (This is the system's entropy over long time scales.)
- **(b)** Compute the entropy again, assuming that the system is in its most likely macrostate. (This is the system's entropy over short time scales, except when there is a large and unlikely fluctuation away from the most likely macrostate.)
- **(c)** Is the issue of time scales really relevant to the entropy of this system?
- **(d)** Suppose that, at a moment when the system is near its most likely macrostate, you suddenly insert a partition between the solids so that they can no longer exchange energy. Now, even over long time scales, the entropy is given by your answer to part (b). Since this number is less than your answer to part (a), you have, in a sense, caused a violation of the second law of thermodynamics. Is this violation significant? Should we lose any sleep over it?

Entropy of an Ideal Gas

The formula for the entropy of a monatomic ideal gas is rather complicated, but extremely useful. If you start with equation 2.40, apply Stirling's approximation, throw away some factors that are merely large, and take the logarithm, you get

$$S = Nk\left[\ln\left(\frac{V}{N}\left(\frac{4\pi mU}{3Nh^2}\right)^{3/2}\right) + \frac{5}{2}\right]. \qquad (2.49)$$

This famous result is known as the **Sackur-Tetrode equation**.

Consider, for instance, a mole of helium at room temperature and atmospheric pressure. The volume is then 0.025 m^3 and the internal energy is $\frac{3}{2}nRT = 3700$ J. Plugging these numbers into the Sackur-Tetrode equation, I find that the argument of the logarithm is 330,000, but the logarithm itself is only 12.7. So the entropy is

$$S = Nk \cdot (15.2) = (9.1 \times 10^{24})k = 126 \text{ J/K}. \tag{2.50}$$

The entropy of an ideal gas depends on its volume, energy, and number of particles. Increasing any of these three variables increases the entropy. The simplest dependence is on the volume; for instance, if the volume changes from V_i to V_f while U and N are held fixed, the entropy changes by

$$\Delta S = Nk \ln \frac{V_f}{V_i} \qquad (U,\, N \text{ fixed}). \tag{2.51}$$

This formula applies, for instance, to the quasistatic isothermal expansion considered in Section 1.5, where the gas pushes on a piston, doing mechanical work, while we simultaneously supply heat from outside to keep the gas at constant temperature. In this case we can think of the entropy increase as being caused by the heat input. Putting heat into a system *always* increases its entropy; in the following chapter I'll discuss in general the relation between entropy and heat.

A very different way of letting a gas expand is shown in Figure 2.14. Initially, the gas is separated by a partition from an evacuated chamber. We then puncture the partition, letting the gas freely expand to fill the whole available space. This process is called **free expansion**. How much work is done during free expansion? None! The gas isn't pushing *on* anything, so it can't do any work. What about heat? Again, none: No heat whatsoever flowed into or out of the gas. Therefore, by the first law of thermodynamics,

$$\Delta U = Q + W = 0 + 0 = 0. \tag{2.52}$$

The energy content of the gas does not change during free expansion, hence formula 2.51 applies. This time, however, the entropy increase was *not* caused by the input of heat; instead we have manufactured *new* entropy, right here on the spot.

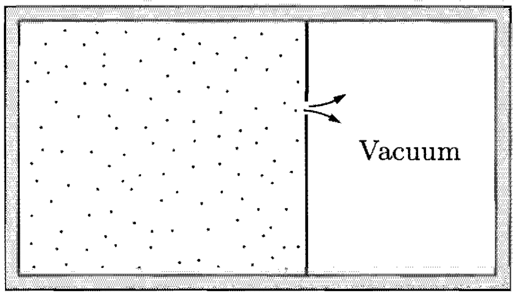


Figure 2.14. Free expansion of a gas into a vacuum. Because the gas neither does work nor absorbs heat, its energy is unchanged. The entropy of the gas increases, however.

Problem 2.31. Fill in the algebraic steps to derive the Sackur-Tetrode equation (2.49).

Problem 2.32. Find an expression for the entropy of the two-dimensional ideal gas considered in Problem 2.26. Express your result in terms of U, A, and N.

Problem 2.33. Use the Sackur-Tetrode equation to calculate the entropy of a mole of argon gas at room temperature and atmospheric pressure. Why is the entropy greater than that of a mole of helium under the same conditions?

Problem 2.34. Show that during the quasistatic isothermal expansion of a monatomic ideal gas, the change in entropy is related to the heat input Q by the simple formula

$$\Delta S = \frac{Q}{T}.$$

In the following chapter I'll prove that this formula is valid for *any* quasistatic process. Show, however, that it is *not* valid for the free expansion process described above.

Problem 2.35. According to the Sackur-Tetrode equation, the entropy of a monatomic ideal gas can become *negative* when its temperature (and hence its energy) is sufficiently low. Of course this is absurd, so the Sackur-Tetrode equation must be invalid at very low temperatures. Suppose you start with a sample of helium at room temperature and atmospheric pressure, then lower the temperature holding the density fixed. Pretend that the helium remains a gas and does not liquefy. Below what temperature would the Sackur-Tetrode equation predict that S is negative? (The behavior of gases at very low temperatures is the main subject of Chapter 7.)

Problem 2.36. For either a monatomic ideal gas or a high-temperature Einstein solid, the entropy is given by Nk times some logarithm. The logarithm is never large, so if all you want is an order-of-magnitude estimate, you can neglect it and just say $S \sim Nk$. That is, the entropy in fundamental units is of the order of the number of particles in the system. This conclusion turns out to be true for most systems (with some important exceptions at low temperatures where the particles are behaving in an orderly way). So just for fun, make a very rough estimate of the entropy of each of the following: this book (a kilogram of carbon compounds); a moose (400 kg of water); the sun (2×10^{30} kg of ionized hydrogen).

Entropy of Mixing

Another way to create entropy is to let two different materials mix with each other. Suppose, for instance, that we start with two different monatomic ideal gases, A and B, each with the same energy, volume, and number of particles. They occupy the two halves of a divided chamber, separated by a partition (see Figure 2.15). If we now remove the partition, the entropy increases. To calculate by how much, we can just treat each gas as a separate system, even after they mix. Since gas A expands to fill twice its initial volume, its entropy increases by

$$\Delta S_A = Nk \ln \frac{V_f}{V_i} = Nk \ln 2, \tag{2.53}$$

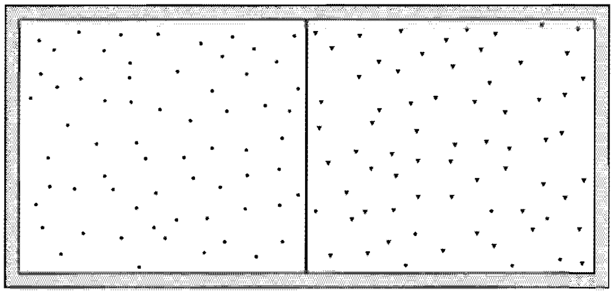


Figure 2.15. Two different gases, separated by a partition. When the partition is removed, each gas expands to fill the whole container, mixing with the other and creating entropy.

while the entropy of gas B increases by the same amount, giving a total increase of

$$\Delta S_{\text{total}} = \Delta S_A + \Delta S_B = 2Nk \ln 2. \tag{2.54}$$

This increase is called the **entropy of mixing.**

It's important to note that this result applies only if the two gases are *different*, like helium and argon. If you start with the *same* gas on both sides, the entropy doesn't increase at all when you remove the partition. (Technically, the total multiplicity does increase, because the distribution of molecules between the two sides can now fluctuate. But the multiplicity increases only by a "large" factor, which has negligible effect on the entropy.)

Let's compare these two situations in a slightly different way. Forget about the partition, and suppose we start with a mole of helium in the chamber. Its total entropy is given by the Sackur-Tetrode equation,

$$S = Nk\left[\ln\left(\frac{V}{N}\left(\frac{4\pi mU}{3Nh^2}\right)^{3/2}\right) + \frac{5}{2}\right]. \tag{2.55}$$

If we now add a mole of argon with the same thermal energy U, the entropy approximately doubles:

$$S_{\text{total}} = S_{\text{helium}} + S_{\text{argon}}. \tag{2.56}$$

(Because the molecular mass enters equation 2.55, the entropy of the argon is actually somewhat greater than the entropy of the helium.) However, if instead we add a second mole of helium, the entropy does *not* double. Look at formula 2.55: If you double the values of both N and U, the ratio U/N inside the logarithm is unchanged, while the N out front becomes $2N$. But there's another N, just inside the logarithm, underneath the V, which also becomes $2N$ and makes the total entropy come out less than you might expect, by a term $2Nk \ln 2$. This "missing" term is precisely the entropy of mixing.

So the difference between adding argon and adding more helium comes from the extra N under the V in the Sackur-Tetrode equation. Where did this N come from? If you look back at the derivation in Section 2.5, you'll see that it came from the $1/N!$ that I slipped into the multiplicity function to account for the fact that molecules in a gas are *indistinguishable* (so interchanging two molecules does not yield a distinct microstate). If I hadn't slipped this factor in, the entropy of a monatomic ideal gas would be

$$S = Nk\left[\ln\left(V\left(\frac{4\pi mU}{3Nh^2}\right)^{3/2}\right) + \frac{3}{2}\right] \qquad \text{(distinguishable molecules).} \qquad (2.57)$$

This formula, if it were correct, would have some rather disturbing consequences. For instance, if you insert a partition into a tank of helium, dividing it in half, this formula predicts that each half would have significantly *less* than half of the original entropy. You could violate the second law of thermodynamics simply by inserting the partition! I don't know an easy way of *proving* that the world isn't like this, but it certainly would be confusing.

This whole issue was first raised by J. Willard Gibbs, and is now known as the **Gibbs paradox**. The best resolution of the paradox is simply to assume that all atoms of a given type are truly indistinguishable. In Chapter 7 we'll see more evidence to support this assumption.

Problem 2.37. Using the same method as in the text, calculate the entropy of mixing for a system of two monatomic ideal gases, A and B, whose relative proportion is arbitrary. Let N be the *total* number of molecules and let x be the fraction of these that are of species B. You should find

$$\Delta S_{\text{mixing}} = -Nk\big[x\ln x + (1-x)\ln(1-x)\big].$$

Check that this expression reduces to the one given in the text when $x = 1/2$.

Problem 2.38. The mixing entropy formula derived in the previous problem actually applies to any ideal gas, and to some dense gases, liquids, and solids as well. For the denser systems, we have to assume that the two types of molecules are the same size and that molecules of different types interact with each other in the same way as molecules of the same type (same forces, etc.). Such a system is called an **ideal mixture**. Explain why, for an ideal mixture, the mixing entropy is given by

$$\Delta S_{\text{mixing}} = k\ln\binom{N}{N_A},$$

where N is the total number of molecules and N_A is the number of molecules of type A. Use Stirling's approximation to show that this expression is the same as the result of the previous problem when both N and N_A are large.

Problem 2.39. Compute the entropy of a mole of helium at room temperature and atmospheric pressure, pretending that all the atoms are distinguishable. Compare to the actual entropy, for indistinguishable atoms, computed in the text.

Reversible and Irreversible Processes

If a physical process increases the total entropy of the universe, that process cannot happen in reverse, since this would violate the second law of thermodynamics. Processes that create new entropy are therefore said to be **irreversible**. By the same token, a process that leaves the total entropy of the universe unchanged would be **reversible**. In practice, no macroscopic process is perfectly reversible, although some processes come close enough for most purposes.

One type of process that creates new entropy is the very sudden expansion of a system, for instance, the free expansion of a gas discussed above. On the other hand, a gradual compression or expansion does not (by itself) change the entropy of a system. In Chapter 3 I'll prove that any reversible volume change must in fact be *quasistatic*, so that $W = -P\Delta V$. (A quasistatic process can still be irreversible, however, if there is also heat flowing in or out or if entropy is being created in some other way.)

It's interesting to think about *why* the slow compression of a gas does not cause its entropy to increase. One way to think about it is to imagine that the molecules in the gas inhabit various quantum-mechanical wavefunctions, each filling the entire box, with discrete (though very closely spaced) energy levels. (See Appendix A for more about the energy levels of particles in a box.) When you compress the gas, each wavefunction gets squeezed, so the energies of all the levels increase, and each molecule's energy increases accordingly. But if the compression is sufficiently slow, molecules will *not* be kicked up into higher energy levels; a molecule that starts in the nth level remains in the nth level (although the energy of that level increases). Thus the number of ways of arranging the molecules among the various energy levels will remain the same, that is, the multiplicity and entropy do not change. On the other hand, if the compression *is* violent enough to kick molecules up into higher levels, then the number of possible arrangements will increase and so will the entropy.

Perhaps the most important type of thermodynamic process is the flow of heat from a hot object to a cold one. We saw in Section 2.3 that this process occurs *because* the total multiplicity of the combined system thereby increases; hence the total entropy increases also, and heat flow is always irreversible. However, we'll see in the next chapter that the increase in entropy becomes negligible in the limit where the temperature difference between the two objects goes to zero. So if you ever hear anyone talking about "reversible heat flow," what they really mean is very slow heat flow, between objects that are at nearly the same temperature. Notice that, in the reversible limit, changing the temperature of one of the objects only infinitesimally can cause the heat to flow in the opposite direction. Similarly, during a quasistatic volume change, an infinitesimal change in the pressure will reverse the direction. In fact, one can *define* a reversible process as one that can be reversed by changing the conditions only infinitesimally.

Most of the processes we observe in life involve large entropy increases and are therefore highly irreversible: sunlight warming the earth, wood burning in the fireplace, metabolism of nutrients in our bodies, mixing ingredients in the kitchen. Because the total entropy of the universe is constantly increasing, and can never decrease, some philosophically inclined physicists have worried that eventually the universe will become a rather boring place: a homogeneous fluid with the maximum possible entropy and no variations in temperature or density anywhere. At the rate we're going, though, this "heat death of the universe" won't occur any time soon; our sun, for instance, should continue to shine brightly for at least another five billion years.*

It may be more fruitful to ask instead about the *beginning* of time. Why did the universe start out in such an improbable, low-entropy state, so that after more than ten billion years it is still so far from equilibrium? Could it have been merely a big coincidence (the biggest of all time)? Or might someone, someday, discover a more satisfying explanation?

Problem 2.40. For each of the following irreversible processes, explain how you can tell that the total entropy of the universe has increased.

- **(a)** Stirring salt into a pot of soup.
- **(b)** Scrambling an egg.
- **(c)** Humpty Dumpty having a great fall.
- **(d)** A wave hitting a sand castle.
- **(e)** Cutting down a tree.
- **(f)** Burning gasoline in an automobile.

Problem 2.41. Describe a few of your favorite, and least favorite, irreversible processes. In each case, explain how you can tell that the entropy of the universe increases.

Problem 2.42. A **black hole** is a region of space where gravity is so strong that nothing, not even light, can escape. Throwing something into a black hole is therefore an irreversible process, at least in the everyday sense of the word. In fact, it is irreversible in the thermodynamic sense as well: Adding mass to a black hole increases the black hole's entropy. It turns out that there's no way to tell (at least from outside) what kind of matter has gone into making a black hole.† Therefore, the entropy of a black hole must be greater than the entropy of any conceivable type of matter that could have been used to create it. Knowing this, it's not hard to estimate the entropy of a black hole.

- **(a)** Use dimensional analysis to show that a black hole of mass M should have a radius of order GM/c^2, where G is Newton's gravitational constant and c is the speed of light. Calculate the approximate radius of a one-solar-mass black hole ($M = 2 \times 10^{30}$ kg).
- **(b)** In the spirit of Problem 2.36, explain why the entropy of a black hole, in fundamental units, should be of the order of the maximum number of particles that could have been used to make it.

---

*For a modern analysis of the long-term prospects for our universe, see Steven Frautschi, "Entropy in an Expanding Universe," *Science* **217**, 593–599 (1982).

†This statement is a slight exaggeration. Electric charge and angular momentum are conserved during black hole formation, and these quantities can still be measured from outside a black hole. In this problem I'm assuming for simplicity that both are zero.

(c) To make a black hole out of the maximum possible number of particles, you should use particles with the lowest possible energy: long-wavelength photons (or other massless particles). But the wavelength can't be any longer than the size of the black hole. By setting the total energy of the photons equal to Mc^2, estimate the maximum number of photons that could be used to make a black hole of mass M. Aside from a factor of $8\pi^2$, your result should agree with the exact formula for the entropy of a black hole, obtained* through a much more difficult calculation:

$$S_{\text{b.h.}} = \frac{8\pi^2 GM^2}{hc}\, k.$$

(d) Calculate the entropy of a one-solar-mass black hole, and comment on the result.

There are 10^{11} *stars in the galaxy. That used to be a huge number. But it's only a hundred billion. It's less than the national deficit! We used to call them astronomical numbers. Now we should call them economical numbers.*

—Richard Feynman, quoted by David Goodstein, *Physics Today* **42**, 73 (February, 1989).

---

*By Stephen Hawking in 1973. To learn more about black hole thermodynamics, see Stephen Hawking, "The Quantum Mechanics of Black Holes," *Scientific American* **236**, 34–40 (January, 1977); Jacob Beckenstein, "Black Hole Thermodynamics," *Physics Today* **33**, 24–31 (January, 1980); and Leonard Susskind, "Black Holes and the Information Paradox," *Scientific American* **276**, 52–57 (April, 1997).