16

Statistical thermodynamics 1: the concepts

Statistical thermodynamics provides the link between the microscopic properties of matter and its bulk properties. Two key ideas are introduced in this chapter. The first is the Boltzmann distribution, which is used to predict the populations of states in systems at thermal equilibrium. In this chapter we see its derivation in terms of the distribution of particles over available states. The derivation leads naturally to the introduction of the partition function, which is the central mathematical concept of this and the next chapter. We see how to interpret the partition function and how to calculate it in a number of simple cases. We then see how to extract thermodynamic information from the partition function. In the final part of the chapter, we generalize the discussion to include systems that are composed of assemblies of interacting particles. Very similar equations are developed to those in the first part of the chapter, but they are much more widely applicable.

The preceding chapters of this part of the text have shown how the energy levels of molecules can be calculated, determined spectroscopically, and related to their structures. The next major step is to see how a knowledge of these energy levels can be used to account for the properties of matter in bulk. To do so, we now introduce the concepts of **statistical thermodynamics**, the link between individual molecular properties and bulk thermodynamic properties.

The crucial step in going from the quantum mechanics of individual molecules to the thermodynamics of bulk samples is to recognize that the latter deals with the *average* behaviour of large numbers of molecules. For example, the pressure of a gas depends on the average force exerted by its molecules, and there is no need to specify which molecules happen to be striking the wall at any instant. Nor is it necessary to consider the fluctuations in the pressure as different numbers of molecules collide with the wall at different moments. The fluctuations in pressure are very small compared with the steady pressure: it is highly improbable that there will be a sudden lull in the number of collisions, or a sudden surge. Fluctuations in other thermodynamic properties also occur, but for large numbers of particles they are negligible compared to the mean values.

This chapter introduces statistical thermodynamics in two stages. The first, the derivation of the Boltzmann distribution for individual particles, is of restricted applicability, but it has the advantage of taking us directly to a result of central importance in a straightforward and elementary way. We can *use* statistical thermodynamics once we have deduced the Boltzmann distribution. Then (in Section 16.5) we extend the arguments to systems composed of interacting particles.

The distribution of molecular states

We consider a closed system composed of N molecules. Although the total energy is constant at E, it is not possible to be definite about how that energy is shared between the molecules. Collisions result in the ceaseless redistribution of energy not only between the molecules but also among their different modes of motion. The closest we can come to a description of the distribution of energy is to report the **population** of a state, the average number of molecules that occupy it, and to say that on average there are n_i molecules in a state of energy ε_i. The populations of the states remain almost constant, but the precise identities of the molecules in each state may change at every collision.

The problem we address in this section is the calculation of the populations of states for any type of molecule in any mode of motion at any temperature. The only restriction is that the molecules should be independent, in the sense that the total energy of the system is a sum of their individual energies. We are discounting (at this stage) the possibility that in a real system a contribution to the total energy may arise from interactions between molecules. We also adopt the **principle of equal *a priori* probabilities**, the assumption that all possibilities for the distribution of energy are equally probable. *A priori* means in this context loosely 'as far as one knows'. We have no reason to presume otherwise than that, for a collection of molecules at thermal equilibrium, vibrational states of a certain energy, for instance, are as likely to be populated as rotational states of the same energy.

One very important conclusion that will emerge from the following analysis is that the populations of states depend on a single parameter, the 'temperature'. That is, statistical thermodynamics provides a molecular justification for the concept of temperature and some insight into this crucially important quantity.

16.1 Configurations and weights

Any individual molecule may exist in states with energies $\varepsilon_0, \varepsilon_1, \ldots$. We shall always take ε_0, the lowest state, as the zero of energy ($\varepsilon_0 = 0$), and measure all other energies relative to that state. To obtain the actual internal energy, U, we may have to add a constant to the calculated energy of the system. For example, if we are considering the vibrational contribution to the internal energy, then we must add the total zero-point energy of any oscillators in the sample.

(a) Instantaneous configurations

At any instant there will be n_0 molecules in the state with energy ε_0, n_1 with ε_1, and so on. The specification of the set of populations $n_0, n_1, \ldots$ in the form $\{n_0, n_1, \ldots\}$ is a statement of the instantaneous **configuration** of the system. The instantaneous configuration fluctuates with time because the populations change. We can picture a large number of different instantaneous configurations. One, for example, might be $\{N,0,0,\ldots\}$, corresponding to every molecule being in its ground state. Another might be $\{N-2,2,0,0,\ldots\}$, in which two molecules are in the first excited state. The latter configuration is intrinsically more likely to be found than the former because it can be achieved in more ways: $\{N,0,0,\ldots\}$ can be achieved in only one way, but $\{N-2,2,0,\ldots\}$ can be achieved in $\frac{1}{2}N(N-1)$ different ways (Fig. 16.1; see *Justification* 16.1). At this stage in the argument, we are ignoring the requirement that the total energy of the system should be constant (the second configuration has a higher energy than the first). The constraint of total energy is imposed later in this section.

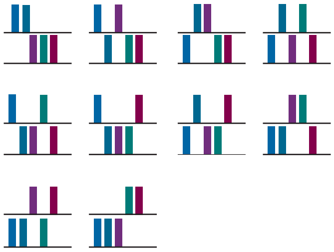

Fig. 16.1 Whereas a configuration $\{5,0,0,\ldots\}$ can be achieved in only one way, a configuration $\{3,2,0,\ldots\}$ can be achieved in the ten different ways shown here, where the tinted blocks represent different molecules.

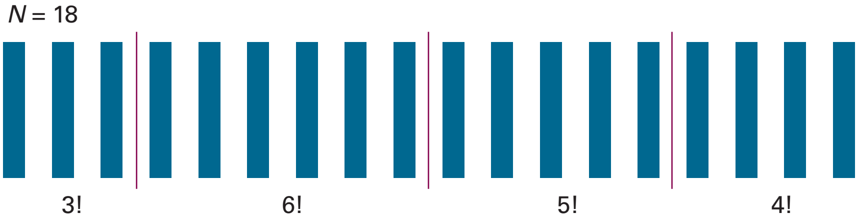

Fig. 16.2 The 18 molecules shown here can be distributed into four receptacles (distinguished by the three vertical lines) in 18! different ways. However, 3! of the selections that put three molecules in the first receptacle are equivalent, 6! that put six molecules into the second receptacle are equivalent, and so on. Hence the number of distinguishable arrangements is 18!/3!6!5!4!.

If, as a result of collisions, the system were to fluctuate between the configurations $\{N,0,0,\ldots\}$ and $\{N-2,2,0,\ldots\}$, it would almost always be found in the second, more likely state (especially if N were large). In other words, a system free to switch between the two configurations would show properties characteristic almost exclusively of the second configuration. A general configuration $\{n_0,n_1,\ldots\}$ can be achieved in W different ways, where W is called the **weight** of the configuration. The weight of the configuration $\{n_0,n_1,\ldots\}$ is given by the expression

$$W = \frac{N!}{n_0!n_1!n_2!\ldots} \tag{16.1}$$

Equation 16.1 is a generalization of the formula $W = \frac{1}{2}N(N-1)$, and reduces to it for the configuration $\{N-2,2,0,\ldots\}$.

Comment 16.1

More formally, W is called the *multinomial coefficient* (see *Appendix* 2). In eqn 16.1, $x!$, x factorial, denotes $x(x-1)(x-2)\ldots 1$, and by definition $0! = 1$.

Justification 16.1 *The weight of a configuration*

First, consider the weight of the configuration $\{N-2,2,0,0,\ldots\}$. One candidate for promotion to an upper state can be selected in N ways. There are $N-1$ candidates for the second choice, so the total number of choices is $N(N-1)$. However, we should not distinguish the choice (Jack, Jill) from the choice (Jill, Jack) because they lead to the same configurations. Therefore, only half the choices lead to distinguishable configurations, and the total number of distinguishable choices is $\frac{1}{2}N(N-1)$.

Now we generalize this remark. Consider the number of ways of distributing N balls into bins. The first ball can be selected in N different ways, the next ball in $N-1$ different ways for the balls remaining, and so on. Therefore, there are $N(N-1)\ldots 1 = N!$ ways of selecting the balls for distribution over the bins. However, if there are n_0 balls in the bin labelled ε_0, there would be $n_0!$ different ways in which the same balls could have been chosen (Fig. 16.2). Similarly, there are $n_1!$ ways in which the n_1 balls in the bin labelled ε_1 can be chosen, and so on. Therefore, the total number of distinguishable ways of distributing the balls so that there are n_0 in bin ε_0, n_1 in bin ε_1, etc. regardless of the order in which the balls were chosen is $N!/n_0!n_1!\ldots$, which is the content of eqn 16.1.

Illustration 16.1 *Calculating the weight of a distribution*

To calculate the number of ways of distributing 20 identical objects with the arrangement 1, 0, 3, 5, 10, 1, we note that the configuration is $\{1,0,3,5,10,1\}$ with $N = 20$; therefore the weight is

$$W = \frac{20!}{1!0!3!5!10!1!} = 9.31 \times 10^{8}$$

Self-test 16.1 Calculate the weight of the configuration in which 20 objects are distributed in the arrangement 0, 1, 5, 0, 8, 0, 3, 2, 0, 1. [4.19×10^{10}]

It will turn out to be more convenient to deal with the natural logarithm of the weight, ln W, rather than with the weight itself. We shall therefore need the expression

$$\begin{aligned}\ln W &= \ln \frac{N!}{n_0!n_1!n_2!\ldots} = \ln N! - \ln(n_0!n_1!n_2!\cdots)\\ &= \ln N! - (\ln n_0! + \ln n_1! + \ln n_2! + \cdots)\\ &= \ln N! - \sum_i \ln n_i!\end{aligned}$$

where in the first line we have used $\ln(x/y) = \ln x - \ln y$ and in the second $\ln xy = \ln x + \ln y$. One reason for introducing ln W is that it is easier to make approximations. In particular, we can simplify the factorials by using **Stirling's approximation** in the form

$$\ln x! \approx x \ln x - x \qquad (16.2)$$

Then the approximate expression for the weight is

$$\ln W = (N \ln N - N) - \sum_i (n_i \ln n_i - n_i) = N \ln N - \sum_i n_i \ln n_i \qquad (16.3)$$

The final form of eqn 16.3 is derived by noting that the sum of n_i is equal to N, so the second and fourth terms in the second expression cancel.

Comment 16.2

A more accurate form of Stirling's approximation is

$$x! \approx (2\pi)^{1/2} x^{x+\frac{1}{2}} e^{-x}$$

and is in error by less than 1 per cent when x is greater than about 10. We deal with far larger values of x, and the simplified version in eqn 16.2 is adequate.

(b) The Boltzmann distribution

We have seen that the configuration $\{N-2,2,0,\ldots\}$ dominates $\{N,0,0,\ldots\}$, and it should be easy to believe that there may be other configurations that have a much greater weight than both. We shall see, in fact, that there is a configuration with so great a weight that it overwhelms all the rest in importance to such an extent that the system will almost always be found in it. The properties of the system will therefore be characteristic of that particular dominating configuration. This dominating configuration can be found by looking for the values of n_i that lead to a maximum value of W. Because W is a function of all the n_i, we can do this search by varying the n_i and looking for the values that correspond to $\mathrm{d}W = 0$ (just as in the search for the maximum of any function), or equivalently a maximum value of ln W. However, there are two difficulties with this procedure.

The first difficulty is that the only permitted configurations are those corresponding to the specified, constant, total energy of the system. This requirement rules out many configurations; $\{N,0,0,\ldots\}$ and $\{N-2,2,0,\ldots\}$, for instance, have different energies, so both cannot occur in the same isolated system. It follows that, in looking for the configuration with the greatest weight, we must ensure that the configuration also satisfies the condition

$$\text{Constant total energy:} \qquad \sum_i n_i \varepsilon_i = E \qquad (16.4)$$

where E is the total energy of the system.

The second constraint is that, because the total number of molecules present is also fixed (at N), we cannot arbitrarily vary all the populations simultaneously. Thus, increasing the population of one state by 1 demands that the population of another state must be reduced by 1. Therefore, the search for the maximum value of W is also subject to the condition

$$\text{Constant total number of molecules:} \qquad \sum_i n_i = N \qquad (16.5)$$

We show in *Further information 16.1* that the populations in the configuration of greatest weight, subject to the two constraints in eqns 16.4 and 16.5, depend on the energy of the state according to the **Boltzmann distribution**:

$$\frac{n_i}{N} = \frac{e^{-\beta\varepsilon_i}}{\sum_i e^{-\beta\varepsilon_i}} \tag{16.6a}$$

where $\varepsilon_0 \le \varepsilon_1 \le \varepsilon_2 \ldots$. Equation 16.6a is the justification of the remark that a single parameter, here denoted β, determines the most probable populations of the states of the system. We shall see in Section 16.3b that

$$\beta = \frac{1}{kT} \tag{16.6b}$$

where T is the thermodynamic temperature and k is Boltzmann's constant. In other words, *the thermodynamic temperature is the unique parameter that governs the most probable populations of states of a system at thermal equilibrium.* In *Further information* 16.3, moreover, we see that β is a more natural measure of temperature than T itself.

16.2 The molecular partition function

From now on we write the Boltzmann distribution as

$$p_i = \frac{e^{-\beta\varepsilon_i}}{q} \tag{16.7}$$

where p_i is the fraction of molecules in the state i, $p_i = n_i/N$, and q is the **molecular partition function**:

$$q = \sum_i e^{-\beta\varepsilon_i} \qquad [16.8]$$

The sum in q is sometimes expressed slightly differently. It may happen that several states have the same energy, and so give the same contribution to the sum. If, for example, g_i states have the same energy ε_i (so the level is g_i-fold degenerate), we could write

$$q = \sum_{\text{levels } i} g_i e^{-\beta\varepsilon_i} \tag{16.9}$$

where the sum is now over energy levels (sets of states with the same energy), not individual states.

Example 16.1 *Writing a partition function*

Write an expression for the partition function of a linear molecule (such as HCl) treated as a rigid rotor.

Method To use eqn 16.9 we need to know (a) the energies of the levels, (b) the degeneracies, the number of states that belong to each level. Whenever calculating a partition function, the energies of the levels are expressed relative to 0 for the state of lowest energy. The energy levels of a rigid linear rotor were derived in Section 13.5c.

Answer From eqn 13.31, the energy levels of a linear rotor are $hcBJ(J+1)$, with $J = 0, 1, 2, \ldots$. The state of lowest energy has zero energy, so no adjustment need be made to the energies given by this expression. Each level consists of $2J+1$ degenerate states. Therefore,

$$q = \sum_{J=0}^{\infty} \overbrace{(2J+1)}^{g_J} e^{\overbrace{-\beta hcBJ(J+1)}^{\varepsilon_J}}$$

The sum can be evaluated numerically by supplying the value of B (from spectroscopy or calculation) and the temperature. For reasons explained in Section 17.2b, this expression applies only to unsymmetrical linear rotors (for instance, HCl, not CO_2).

Self-test 16.2 Write the partition function for a two-level system, the lower state (at energy 0) being nondegenerate, and the upper state (at an energy ε) doubly degenerate. [$q = 1 + 2e^{-\beta\varepsilon}$]

(a) An interpretation of the partition function

Some insight into the significance of a partition function can be obtained by considering how q depends on the temperature. When T is close to zero, the parameter $\beta = 1/kT$ is close to infinity. Then every term except one in the sum defining q is zero because each one has the form e^{-x} with $x \to \infty$. The exception is the term with $\varepsilon_0 \equiv 0$ (or the g_0 terms at zero energy if the ground state is g_0-fold degenerate), because then $\varepsilon_0/kT \equiv 0$ whatever the temperature, including zero. As there is only one surviving term when $T = 0$, and its value is g_0, it follows that

$$\lim_{T\to 0} q = g_0 \qquad (16.10)$$

That is, at $T = 0$, the partition function is equal to the degeneracy of the ground state.

Now consider the case when T is so high that for each term in the sum $\varepsilon_j/kT \approx 0$. Because $e^{-x} = 1$ when $x = 0$, each term in the sum now contributes 1. It follows that the sum is equal to the number of molecular states, which in general is infinite:

$$\lim_{T\to \infty} q = \infty \qquad (16.11)$$

In some idealized cases, the molecule may have only a finite number of states; then the upper limit of q is equal to the number of states. For example, if we were considering only the spin energy levels of a radical in a magnetic field, then there would be only two states ($m_s = \pm\frac{1}{2}$). The partition function for such a system can therefore be expected to rise towards 2 as T is increased towards infinity.

We see that *the molecular partition function gives an indication of the number of states that are thermally accessible to a molecule at the temperature of the system*. At $T = 0$, only the ground level is accessible and $q = g_0$. At very high temperatures, virtually all states are accessible, and q is correspondingly large.

Example 16.2 *Evaluating the partition function for a uniform ladder of energy levels*

Evaluate the partition function for a molecule with an infinite number of equally spaced nondegenerate energy levels (Fig. 16.3). These levels can be thought of as the vibrational energy levels of a diatomic molecule in the harmonic approximation.

Method We expect the partition function to increase from 1 at $T = 0$ and approach infinity as T to ∞. To evaluate eqn 16.8 explicitly, note that

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

Answer If the separation of neighbouring levels is ε, the partition function is

$$q = 1 + e^{-\beta\varepsilon} + e^{-2\beta\varepsilon} + \cdots = 1 + e^{-\beta\varepsilon} + (e^{-\beta\varepsilon})^2 + \cdots = \frac{1}{1 - e^{-\beta\varepsilon}}$$

This expression is plotted in Fig. 16.4: notice that, as anticipated, q rises from 1 to infinity as the temperature is raised.

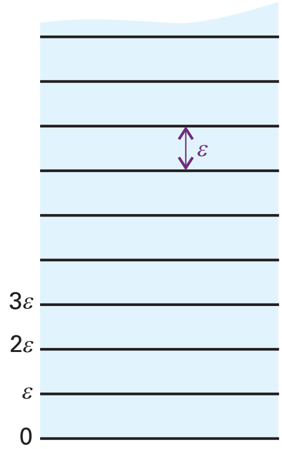

Fig. 16.3 The equally spaced infinite array of energy levels used in the calculation of the partition function. A harmonic oscillator has the same spectrum of levels.

Comment 16.3

The sum of the infinite series $S = 1 + x + x^2 + \cdots$ is obtained by multiplying both sides by x, which gives $xS = x + x^2 + x^3 + \cdots = S - 1$ and hence $S = 1/(1 - x)$.

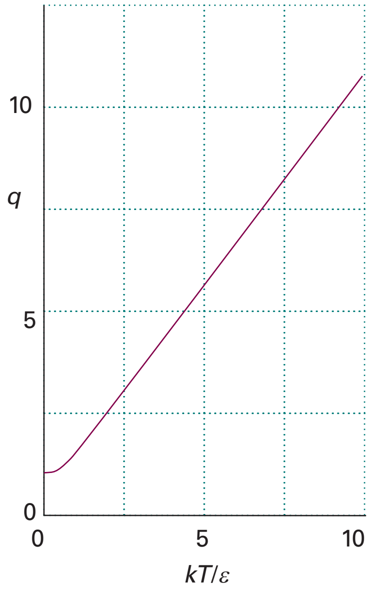

Fig. 16.4 The partition function for the system shown in Fig.16.3 (a harmonic oscillator) as a function of temperature.

Exploration Plot the partition function of a harmonic oscillator against temperature for several values of the energy separation ε. How does q vary with temperature when T is high, in the sense that $kT \gg \varepsilon$ (or $\beta\varepsilon \ll 1$)?

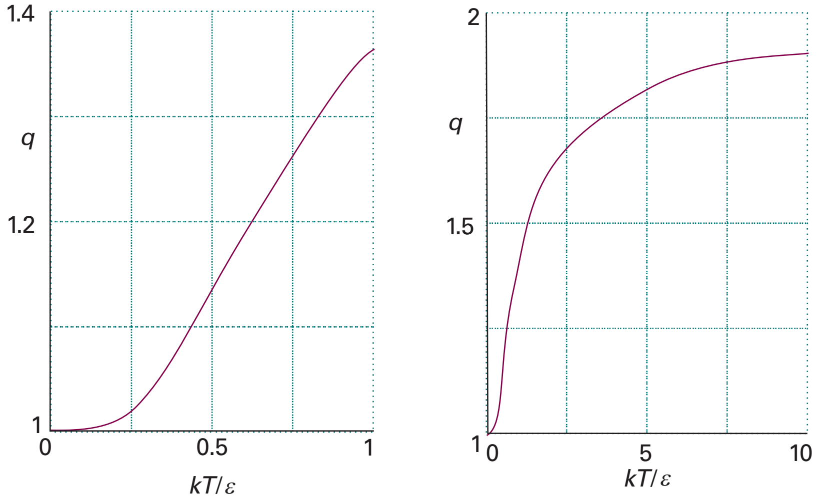

Fig. 16.5 The partition function for a two-level system as a function of temperature. The two graphs differ in the scale of the temperature axis to show the approach to 1 as $T \to 0$ and the slow approach to 2 as $T \to \infty$.

Exploration Consider a three-level system with levels 0, ε, and 2ε. Plot the partition function against kT/ε.

Self-test 16.3 Find and plot an expression for the partition function of a system with one state at zero energy and another state at the energy ε.

[$q = 1 + e^{-\beta\varepsilon}$, Fig. 16.5]

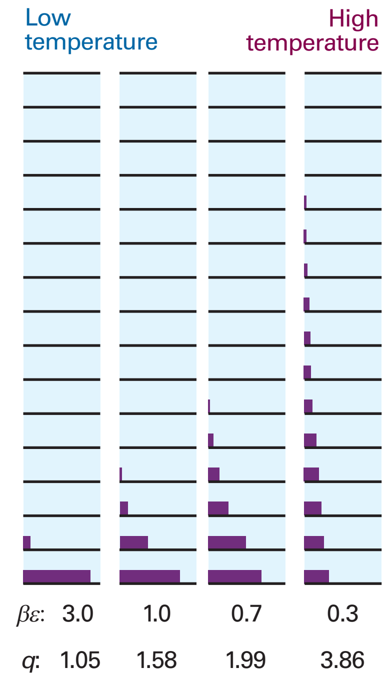

Fig. 16.6 The populations of the energy levels of the system shown in Fig.16.3 at different temperatures, and the corresponding values of the partition function calculated in Example 16.2. Note that $\beta = 1/kT$.

Exploration To visualize the content of Fig. 16.6 in a different way, plot the functions p_0, p_1, p_2, and p_3 against kT/ε.

It follows from eqn 16.8 and the expression for q derived in Example 16.2 for a uniform ladder of states of spacing ε,

$$q = \frac{1}{1 - e^{-\beta\varepsilon}} \qquad (16.12)$$

that the fraction of molecules in the state with energy ε_i is

$$p_i = \frac{e^{-\beta\varepsilon_i}}{q} = (1 - e^{-\beta\varepsilon})e^{-\beta\varepsilon_i} \qquad (16.13)$$

Figure 16.6 shows how p_i varies with temperature. At very low temperatures, where q is close to 1, only the lowest state is significantly populated. As the temperature is raised, the population breaks out of the lowest state, and the upper states become progressively more highly populated. At the same time, the partition function rises from 1 and its value gives an indication of the range of states populated. The name 'partition function' reflects the sense in which q measures how the total number of molecules is distributed—partitioned—over the available states.

The corresponding expressions for a two-level system derived in Self-test 16.3 are

$$p_0 = \frac{1}{1 + e^{-\beta\varepsilon}} \qquad p_1 = \frac{e^{-\beta\varepsilon}}{1 + e^{-\beta\varepsilon}} \qquad (16.14)$$

These functions are plotted in Fig. 16.7. Notice how the populations tend towards equality ($p_0 = \tfrac{1}{2}$, $p_1 = \tfrac{1}{2}$) as $T \to \infty$. A common error is to suppose that all the molecules in the system will be found in the upper energy state when $T = \infty$; however, we see

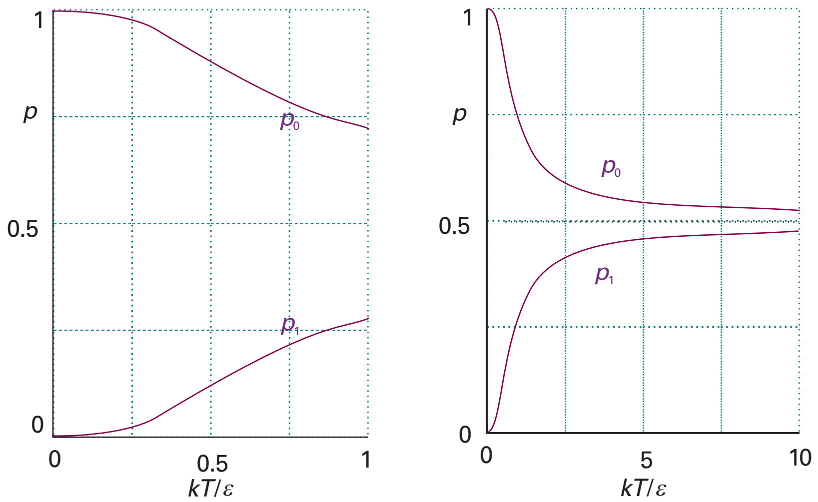

Fig. 16.7 The fraction of populations of the two states of a two-level system as a function of temperature (eqn 16.14). Note that, as the temperature approaches infinity, the populations of the two states become equal (and the fractions both approach 0.5).

Exploration Consider a three-level system with levels 0, ε, and 2ε. Plot the functions p_0, p_1, and p_2 against kT/ε.

from eqn 16.14 that, as $T \rightarrow \infty$, the populations of states become equal. The same conclusion is true of multi-level systems too: as $T \rightarrow \infty$, all states become equally populated.

Example 16.3 *Using the partition function to calculate a population*

Calculate the proportion of I_2 molecules in their ground, first excited, and second excited vibrational states at 25°C. The vibrational wavenumber is 214.6 cm^{-1}.

Method Vibrational energy levels have a constant separation (in the harmonic approximation, Section 13.9), so the partition function is given by eqn 16.12 and the populations by eqn 16.13. To use the latter equation, we identify the index i with the quantum number υ, and calculate p_υ for $\upsilon = 0$, 1, and 2. At 298.15 K, $kT/hc = 207.226$ cm^{-1}.

Answer First, we note that

$$\beta\varepsilon = \frac{hc\tilde{\nu}}{kT} = \frac{214.6\ \text{cm}^{-1}}{207.226\ \text{cm}^{-1}} = 1.036$$

Then it follows from eqn 16.13 that the populations are

$$p_\upsilon = (1 - \mathrm{e}^{-\beta\varepsilon})\mathrm{e}^{-\upsilon\beta\varepsilon} = 0.645\mathrm{e}^{-1.036\upsilon}$$

Therefore, $p_0 = 0.645$, $p_1 = 0.229$, $p_2 = 0.081$. The I—I bond is not stiff and the atoms are heavy: as a result, the vibrational energy separations are small and at room temperature several vibrational levels are significantly populated. The value of the partition function, $q = 1.55$, reflects this small but significant spread of populations.

Self-test 16.4 At what temperature would the $\upsilon = 1$ level of I_2 have (a) half the population of the ground state, (b) the same population as the ground state?

[(a) 445 K, (b) infinite]

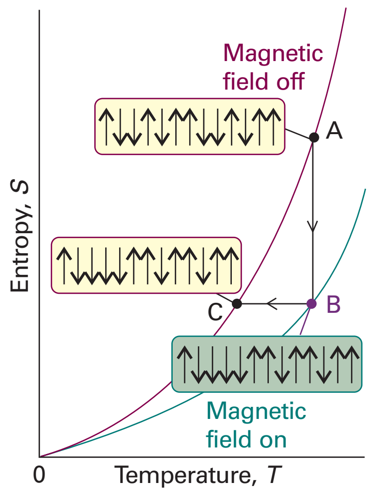

Fig. 16.8 The technique of adiabatic demagnetization is used to attain very low temperatures. The upper curve shows that variation of the entropy of a paramagnetic system in the absence of an applied field. The lower curve shows that variation in entropy when a field is applied and has made the electron magnets more orderly. The isothermal magnetization step is from A to B; the adiabatic demagnetization step (at constant entropy) is from B to C.

It follows from our discussion of the partition function that to reach low temperatures it is necessary to devise strategies that populate the low energy levels of a system at the expense of high energy levels. Common methods used to reach very low temperatures include **optical trapping** and **adiabatic demagnetization**. In optical trapping, atoms in the gas phase are cooled by inelastic collisions with photons from intense laser beams, which act as walls of a very small container. Adiabatic demagnetization is based on the fact that, in the absence of a magnetic field, the unpaired electrons of a paramagnetic material are orientated at random, but in the presence of a magnetic field there are more β spins ($m_s = -\frac{1}{2}$) than α spins ($m_s = +\frac{1}{2}$). In thermodynamic terms, the application of a magnetic field lowers the entropy of a sample and, at a given temperature, the entropy of a sample is lower when the field is on than when it is off. Even lower temperatures can be reached if nuclear spins (which also behave like small magnets) are used instead of electron spins in the technique of **adiabatic nuclear demagnetization**, which has been used to cool a sample of silver to about 280 pK. In certain circumstances it is possible to achieve negative temperatures, and the equations derived later in this chapter can be extended to $T < 0$ with interesting consequences (see *Further information* 16.3).

Illustration 16.2 *Cooling a sample by adiabatic demagnetization*

Consider the situation summarized by Fig. 16.8. A sample of paramagnetic material, such as a *d*- or *f*-metal complex with several unpaired electrons, is cooled to about 1 K by using helium. The sample is then exposed to a strong magnetic field while it is surrounded by helium, which provides thermal contact with the cold reservoir. This magnetization step is isothermal, and energy leaves the system as heat while the electron spins adopt the lower energy state (AB in the illustration). Thermal contact between the sample and the surroundings is now broken by pumping away the helium and the magnetic field is reduced to zero. This step is adiabatic and effectively reversible, so the state of the sample changes from B to C. At the end of this step the sample is the same as it was at A except that it now has a lower entropy. That lower entropy in the absence of a magnetic field corresponds to a lower temperature. That is, adiabatic demagnetization has cooled the sample.

(b) Approximations and factorizations

In general, exact analytical expressions for partition functions cannot be obtained. However, closed approximate expressions can often be found and prove to be very important in a number of chemical and biochemical applications (*Impact* 16.1). For instance, the expression for the partition function for a particle of mass m free to move in a one-dimensional container of length X can be evaluated by making use of the fact that the separation of energy levels is very small and that large numbers of states are accessible at normal temperatures. As shown in the *Justification* below, in this case

$$q_X = \left(\frac{2\pi m}{h^2\beta}\right)^{1/2} X \tag{16.15}$$

This expression shows that the partition function for translational motion increases with the length of the box and the mass of the particle, for in each case the separation of the energy levels becomes smaller and more levels become thermally accessible. For a given mass and length of the box, the partition function also increases with increasing temperature (decreasing β), because more states become accessible.

Justification 16.2 *The partition function for a particle in a one-dimensional box*

The energy levels of a molecule of mass m in a container of length X are given by eqn 9.4a with $L = X$:

$$E_n = \frac{n^2h^2}{8mX^2} \qquad n = 1, 2, \ldots$$

The lowest level ($n = 1$) has energy $h^2/8mX^2$, so the energies relative to that level are

$$\varepsilon_n = (n^2 - 1)\varepsilon \qquad \varepsilon = h^2/8mX^2$$

The sum to evaluate is therefore

$$q_X = \sum_{n=1}^{\infty} \mathrm{e}^{-(n^2-1)\beta\varepsilon}$$

The translational energy levels are very close together in a container the size of a typical laboratory vessel; therefore, the sum can be approximated by an integral:

$$q_X = \int_1^{\infty} \mathrm{e}^{-(n^2-1)\beta\varepsilon}\mathrm{d}n \approx \int_0^{\infty} \mathrm{e}^{-n^2\beta\varepsilon}\mathrm{d}n$$

The extension of the lower limit to $n = 0$ and the replacement of $n^2 - 1$ by n^2 introduces negligible error but turns the integral into standard form. We make the substitution $x^2 = n^2\beta\varepsilon$, implying $\mathrm{d}n = \mathrm{d}x/(\beta\varepsilon)^{1/2}$, and therefore that

$$q_X = \left(\frac{1}{\beta\varepsilon}\right)^{1/2} \overbrace{\int_0^{\infty} \mathrm{e}^{-x^2}\mathrm{d}x}^{\pi^{1/2}/2} = \left(\frac{1}{\beta\varepsilon}\right)^{1/2}\left(\frac{\pi^{1/2}}{2}\right) = \left(\frac{2\pi m}{h^2\beta}\right)^{1/2} X$$

Another useful feature of partition functions is used to derive expressions when the energy of a molecule arises from several different, independent sources: if the energy is a sum of contributions from independent modes of motion, then the partition function is a product of partition functions for each mode of motion. For instance, suppose the molecule we are considering is free to move in three dimensions. We take the length of the container in the y-direction to be Y and that in the z-direction to be Z. The total energy of a molecule ε is the sum of its translational energies in all three directions:

$$\varepsilon_{n_1n_2n_3} = \varepsilon_{n_1}^{(X)} + \varepsilon_{n_2}^{(Y)} + \varepsilon_{n_3}^{(Z)} \tag{16.16}$$

where n_1, n_2, and n_3 are the quantum numbers for motion in the x-, y-, and z-directions, respectively. Therefore, because $\mathrm{e}^{a+b+c} = \mathrm{e}^a\mathrm{e}^b\mathrm{e}^c$, the partition function factorizes as follows:

$$\begin{aligned} q &= \sum_{\text{all } n} \mathrm{e}^{-\beta\varepsilon_{n_1}^{(X)} - \beta\varepsilon_{n_2}^{(Y)} - \beta\varepsilon_{n_3}^{(Z)}} = \sum_{\text{all } n} \mathrm{e}^{-\beta\varepsilon_{n_1}^{(X)}}\mathrm{e}^{-\beta\varepsilon_{n_2}^{(Y)}}\mathrm{e}^{-\beta\varepsilon_{n_3}^{(Z)}} \\ &= \left(\sum_{n_1} \mathrm{e}^{-\beta\varepsilon_{n_1}^{(X)}}\right)\left(\sum_{n_2} \mathrm{e}^{-\beta\varepsilon_{n_2}^{(Y)}}\right)\left(\sum_{n_3} \mathrm{e}^{-\beta\varepsilon_{n_3}^{(Z)}}\right) \qquad (16.17) \\ &= q_X q_Y q_Z \end{aligned}$$

It is generally true that, if the energy of a molecule can be written as the sum of independent terms, then the partition function is the corresponding product of individual contributions.

Equation 16.15 gives the partition function for translational motion in the x-direction. The only change for the other two directions is to replace the length X by the lengths Y or Z. Hence the partition function for motion in three dimensions is

$$q = \left(\frac{2\pi m}{h^2\beta}\right)^{3/2} XYZ \qquad (16.18)$$

The product of lengths XYZ is the volume, V, of the container, so we can write

$$q = \frac{V}{\Lambda^3} \qquad \Lambda = h\left(\frac{\beta}{2\pi m}\right)^{1/2} = \frac{h}{(2\pi mkT)^{1/2}} \qquad (16.19)$$

The quantity Λ has the dimensions of length and is called the **thermal wavelength** (sometimes the *thermal de Broglie wavelength*) of the molecule. The thermal wavelength decreases with increasing mass and temperature. As in the one-dimensional case, the partition function increases with the mass of the particle (as $m^{3/2}$) and the volume of the container (as V); for a given mass and volume, the partition function increases with temperature (as $T^{3/2}$).

Illustration 16.3 *Calculating the translational partition function*

To calculate the translational partition function of an H_2 molecule confined to a 100 cm^3 vessel at 25°C we use $m = 2.016$ u; then

$$\Lambda = \frac{6.626\times10^{-34}\ \text{J s}}{\{2\pi\times(2.016\times1.6605\times10^{-27}\ \text{kg})\times(1.38\times10^{-23}\ \text{J K}^{-1})\times(298\ \text{K})\}^{1/2}}$$

$$= 7.12\times10^{-11}\ \text{m}$$

where we have used $1\ \text{J} = 1\ \text{kg m}^2\ \text{s}^{-2}$. Therefore,

$$q = \frac{1.00\times10^{-4}\ \text{m}^3}{(7.12\times10^{-11}\ \text{m})^3} = 2.77\times10^{26}$$

About 10^{26} quantum states are thermally accessible, even at room temperature and for this light molecule. Many states are occupied if the thermal wavelength (which in this case is 71.2 pm) is small compared with the linear dimensions of the container.

Self-test 16.5 Calculate the translational partition function for a D_2 molecule under the same conditions. [$q = 7.8\times10^{26}$, $2^{3/2}$ times larger]

The validity of the approximations that led to eqn 16.19 can be expressed in terms of the average separation of the particles in the container, d. We do not have to worry about the role of the Pauli principle on the occupation of states if there are many states available for each molecule. Because q is the total number of accessible states, the average number of states per molecule is q/N. For this quantity to be large, we require $V/N\Lambda^3 \gg 1$. However, V/N is the volume occupied by a single particle, and therefore the average separation of the particles is $d = (V/N)^{1/3}$. The condition for there being many states available per molecule is therefore $d^3/\Lambda^3 \gg 1$, and therefore $d \gg \Lambda$. That is, for eqn 16.19 to be valid, *the average separation of the particles must be much greater than their thermal wavelength*. For H_2 molecules at 1 bar and 298 K, the average separation is 3 nm, which is significantly larger than their thermal wavelength (71.2 pm, *Illustration* 16.3).

IMPACT ON BIOCHEMISTRY

I16.1 The helix–coil transition in polypeptides

Proteins are polymers that attain well defined three-dimensional structures both in solution and in biological cells. They are *polypeptides* formed from different amino acids strung together by the *peptide link*, —CONH—. Hydrogen bonds between amino acids of a polypeptide give rise to stable helical or sheet structures, which may collapse into a random coil when certain conditions are changed. The unwinding of a helix into a random coil is a *cooperative transition*, in which the polymer becomes increasingly more susceptible to structural changes once the process has begun. We examine here a model grounded in the principles of statistical thermodynamics that accounts for the cooperativity of the helix–coil transition in polypeptides.

To calculate the fraction of polypeptide molecules present as helix or coil we need to set up the partition function for the various states of the molecule. To illustrate the approach, consider a short polypeptide with four amino acid residues, each labelled *h* if it contributes to a helical region and *c* if it contributes to a random coil region. We suppose that conformations *hhhh* and *cccc* contribute terms q_0 and q_4, respectively, to the partition function q. Then we assume that each of the four conformations with one *c* amino acid (such as *hchh*) contributes q_1. Similarly, each of the six states with two *c* amino acids contributes a term q_2, and each of the four states with three *c* amino acids contributes a term q_3. The partition function is then

$$q = q_0 + 4q_1 + 6q_2 + 4q_3 + q_4 = q_0\left(1 + \frac{4q_1}{q_0} + \frac{6q_2}{q_0} + \frac{4q_3}{q_0} + \frac{q_4}{q_0}\right)$$

We shall now suppose that each partition function differs from q_0 only by the energy of each conformation relative to *hhhh*, and write

$$\frac{q_i}{q_0} = \mathrm{e}^{-(\varepsilon_i - \varepsilon_0)/kT}$$

Next, we suppose that the conformational transformations are non-cooperative, in the sense that the energy associated with changing one *h* amino acid into one *c* amino acid has the same value regardless of how many *h* or *c* amino acid residues are in the reactant or product state and regardless of where in the chain the conversion occurs. That is, we suppose that the difference in energy between $c^i h^{4-i}$ and $c^{i+1} h^{3-i}$ has the same value γ for all i. This assumption implies that $\varepsilon_i - \varepsilon_0 = i\gamma$ and therefore that

$$q = q_0(1 + 4s + 6s^2 + 4s^3 + s^4) \qquad s = \mathrm{e}^{-\Gamma/RT} \tag{16.20}$$

where $\Gamma = N_A\gamma$ and s is called the *stability parameter*. The term in parentheses has the form of the binomial expansion of $(1 + s)^4$.

$$\frac{q}{q_0} = \sum_{i=0}^{4} C(4,i)s^i \quad \text{with} \quad C(4,i) = \frac{4!}{(4-i)!i!} \tag{16.21}$$

which we interpret as the number of ways in which a state with i *c* amino acids can be formed.

The extension of this treatment to take into account a longer chain of residues is now straightforward: we simply replace the upper limit of 4 in the sum by n:

$$\frac{q}{q_0} = \sum_{i=0}^{n} C(n,i)s^i \tag{16.22}$$

A cooperative transformation is more difficult to accommodate, and depends on building a model of how neighbours facilitate each other's conformational change. In

Comment 16.4

The binomial expansion of $(1 + x)^n$ is

$$(1 + x)^n = \sum_{i=0}^{n} C(n,i)x^i,$$

$$\text{with } C(n,i) = \frac{n!}{(n-i)!i!}$$

the simple *zipper model*, conversion from h to c is allowed only if a residue adjacent to the one undergoing the conversion is already a c residue. Thus, the zipper model allows a transition of the type . . .*hhhch*. . . → . . .*hhhcc*. . . , but not a transition of the type . . .*hhhch*. . . → . . .*hchch*. . . . The only exception to this rule is, of course, the very first conversion from h to c in a fully helical chain. Cooperativity is included in the zipper model by assuming that the first conversion from h to c, called the *nucleation step*, is less favourable than the remaining conversions and replacing s for that step by σs, where $\sigma \ll 1$. Each subsequent step is called a *propagation step* and has a stability parameter s. In Problem 16.24, you are invited to show that the partition function is:

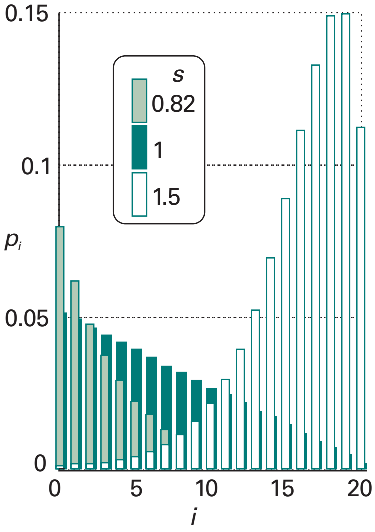

Fig. 16.9 The distribution of p_i, the fraction of molecules that has a number i of c amino acids for $s = 0.8$ ($\langle i \rangle = 1.1$), 1.0 ($\langle i \rangle = 3.8$), and 1.5 ($\langle i \rangle = 15.9$), with $\sigma = 5.0 \times 10^{-3}$.

$$q = 1 + \sum_{i=1}^{n} Z(n,i)\sigma s^i \tag{16.23}$$

where $Z(n,i)$ is the number of ways in which a state with a number i of c amino acids can be formed under the strictures of the zipper model. Because $Z(n,i) = n - i + 1$ (see Problem 16.24),

$$q = 1 + \sigma(n+1)\sum_{i=1}^{n} s^i - \sigma\sum_{i=1}^{n} is^i \tag{16.24}$$

After evaluating both geometric series by using the two relations

$$\sum_{i=1}^{n} x^i = \frac{x^{n+1} - x}{x - 1} \qquad \sum_{i=1}^{n} ix^i = \frac{x}{(x-1)^2}[nx^{n+1} - (n+1)x^n + 1]$$

we find

$$q = 1 + \frac{\sigma s[s^{n+1} - (n+1)s^n + 1]}{(s-1)^2}$$

The fraction $p_i = q_i/q$ of molecules that has a number i of c amino acids is $p_i = [(n - i + 1)\sigma s^i]/q$ and the mean value of i is then $\langle i \rangle = \sum_i ip_i$. Figure 16.9 shows the distribution of p_i for various values of s with $\sigma = 5.0 \times 10^{-3}$. We see that most of the polypeptide chains remain largely helical when $s < 1$ and that most of the chains exist largely as random coils when $s > 1$. When $s = 1$, there is a more widespread distribution of length of random coil segments. Because the *degree of conversion*, θ, of a polypeptide with n amino acids to a random coil is defined as $\theta = \langle i \rangle/n$, it is possible to show (see Problem 16.24) that

$$\theta = \frac{1}{n}\frac{\mathrm{d}}{\mathrm{d}(\ln s)}\ln q \tag{16.25}$$

This is a general result that applies to any model of the helix–coil transition in which the partition function q is expressed as a function of the stability parameter s.

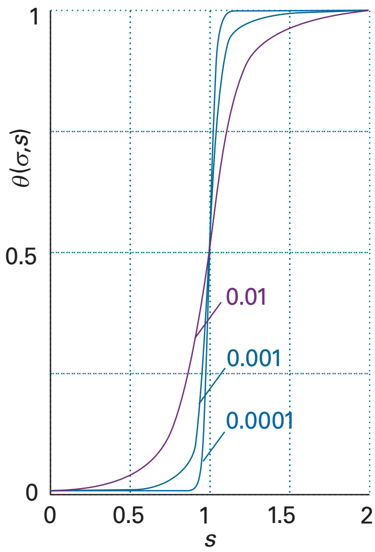

Fig. 16.10 Plots of the degree of conversion θ, against s for several values of σ. The curves show the sigmoidal shape characteristics of cooperative behaviour.

A more sophisticated model for the helix–coil transition must allow for helical segments to form in different regions of a long polypeptide chain, with the nascent helices being separated by shrinking coil segments. Calculations based on this more complete *Zimm–Bragg model* give

$$\theta = \tfrac{1}{2}\left(1 + \frac{(s-1) + 2\sigma}{[(s-1)^2 + 4s\sigma]^{1/2}}\right) \tag{16.26}$$

Figure 16.10 shows plots of θ against s for several values of σ. The curves show the sigmoidal shape characteristic of cooperative behaviour. There is a sudden surge of transition to a random coil as s passes through 1 and, the smaller the parameter σ, the greater the sharpness and hence the greater the cooperativity of the transition. That is, the harder it is to get coil formation started, the sharper the transition from helix to coil.

The internal energy and the entropy

The importance of the molecular partition function is that it contains all the information needed to calculate the thermodynamic properties of a system of independent particles. In this respect, q plays a role in statistical thermodynamics very similar to that played by the wavefunction in quantum mechanics: q is a kind of thermal wavefunction.

16.3 The internal energy

We shall begin to unfold the importance of q by showing how to derive an expression for the internal energy of the system.

(a) The relation between U and q

The total energy of the system relative to the energy of the lowest state is

$$E=\sum_i n_i\varepsilon_i \tag{16.27}$$

Because the most probable configuration is so strongly dominating, we can use the Boltzmann distribution for the populations and write

$$E=\frac{N}{q}\sum_i \varepsilon_i \mathrm{e}^{-\beta\varepsilon_i} \tag{16.28}$$

To manipulate this expression into a form involving only q we note that

$$\varepsilon_i \mathrm{e}^{-\beta\varepsilon_i}=-\frac{\mathrm{d}}{\mathrm{d}\beta}\mathrm{e}^{-\beta\varepsilon_i}$$

It follows that

$$E=-\frac{N}{q}\sum_i \frac{\mathrm{d}}{\mathrm{d}\beta}\mathrm{e}^{-\beta\varepsilon_i}=-\frac{N}{q}\frac{\mathrm{d}}{\mathrm{d}\beta}\sum_i \mathrm{e}^{-\beta\varepsilon_i}=-\frac{N}{q}\frac{\mathrm{d}q}{\mathrm{d}\beta} \tag{16.29}$$

Illustration 16.4 *The energy of a two-level system*

From the two-level partition function $q = 1 + \mathrm{e}^{-\beta\varepsilon}$, we can deduce that the total energy of N two-level systems is

$$E=-\left(\frac{N}{1+\mathrm{e}^{-\beta\varepsilon}}\right)\frac{\mathrm{d}}{\mathrm{d}\beta}(1+\mathrm{e}^{-\beta\varepsilon})=\frac{N\varepsilon\mathrm{e}^{-\beta\varepsilon}}{1+\mathrm{e}^{-\beta\varepsilon}}=\frac{N\varepsilon}{1+\mathrm{e}^{\beta\varepsilon}}$$

This function is plotted in Fig. 16.11. Notice how the energy is zero at $T=0$, when only the lower state (at the zero of energy) is occupied, and rises to $\frac{1}{2}N\varepsilon$ as $T\rightarrow\infty$, when the two levels become equally populated.

There are several points in relation to eqn 16.29 that need to be made. Because $\varepsilon_0 = 0$ (remember that we measure all energies from the lowest available level), E should be interpreted as the value of the internal energy relative to its value at $T=0$, $U(0)$. Therefore, to obtain the conventional internal energy U, we must add the internal energy at $T=0$:

$$U=U(0)+E \tag{16.30}$$

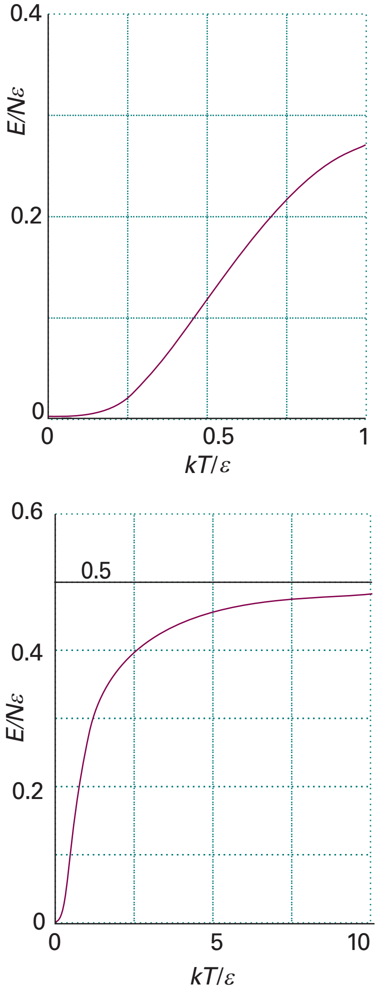

Fig. 16.11 The total energy of a two-level system (expressed as a multiple of $N\varepsilon$) as a function of temperature, on two temperature scales. The graph at the top shows the slow rise away from zero energy at low temperatures; the slope of the graph at $T=0$ is 0 (that is, the heat capacity is zero at $T=0$). The graph below shows the slow rise to 0.5 as $T\rightarrow\infty$ as both states become equally populated (see Fig. 16.7).

Exploration Draw graphs similar to those in Fig. 16.11 for a three-level system with levels 0, ε, and 2ε.

Secondly, because the partition function may depend on variables other than the temperature (for example, the volume), the derivative with respect to β in eqn 16.29 is actually a *partial* derivative with these other variables held constant. The complete expression relating the molecular partition function to the thermodynamic internal energy of a system of independent molecules is therefore

$$U = U(0) - \frac{N}{q}\left(\frac{\partial q}{\partial \beta}\right)_V \tag{16.31a}$$

An equivalent form is obtained by noting that $\mathrm{d}x/x = \mathrm{d}\ln x$:

$$U = U(0) - N\left(\frac{\partial \ln q}{\partial \beta}\right)_V \tag{16.31b}$$

These two equations confirm that we need know only the partition function (as a function of temperature) to calculate the internal energy relative to its value at $T = 0$.

(b) The value of β

We now confirm that the parameter β, which we have anticipated is equal to $1/kT$, does indeed have that value. To do so, we compare the equipartition expression for the internal energy of a monatomic perfect gas, which from *Molecular interpretation* 2.2 we know to be

$$U = U(0) + \tfrac{3}{2}nRT \tag{16.32a}$$

with the value calculated from the translational partition function (see the following *Justification*), which is

$$U = U(0) + \frac{3N}{2\beta} \tag{16.32b}$$

It follows by comparing these two expressions that

$$\beta = \frac{N}{nRT} = \frac{nN_A}{nN_AkT} = \frac{1}{kT} \tag{16.33}$$

(We have used $N = nN_A$, where n is the amount of gas molecules, N_A is Avogadro's constant, and $R = N_Ak$.) Although we have proved that $\beta = 1/kT$ by examining a very specific example, the translational motion of a perfect monatomic gas, the result is general (see *Example* 17.1 and *Further reading*).

Justification 16.3 *The internal energy of a perfect gas*

To use eqn 16.31, we introduce the translational partition function from eqn 16.19:

$$\left(\frac{\partial q}{\partial \beta}\right)_V = \left(\frac{\partial}{\partial \beta}\frac{V}{\Lambda^3}\right)_V = V\frac{\mathrm{d}}{\mathrm{d}\beta}\frac{1}{\Lambda^3} = -3\frac{V}{\Lambda^4}\frac{\mathrm{d}\Lambda}{\mathrm{d}\beta}$$

Then we note from the formula for Λ in eqn 16.19 that

$$\frac{\mathrm{d}\Lambda}{\mathrm{d}\beta} = \frac{\mathrm{d}}{\mathrm{d}\beta}\left\{\frac{h\beta^{1/2}}{(2\pi m)^{1/2}}\right\} = \frac{1}{2\beta^{1/2}} \times \frac{h}{(2\pi m)^{1/2}} = \frac{\Lambda}{2\beta}$$

and so obtain

$$\left(\frac{\partial q}{\partial \beta}\right)_V = -\frac{3V}{2\beta\Lambda^3}$$

Then, by eqn 16.31a,

$$U = U(0) - N\left(\frac{\Lambda^3}{V}\right)\left(-\frac{3V}{2\beta\Lambda^3}\right) = U(0) + \frac{3N}{2\beta}$$

as in eqn 16.32b.

16.4 The statistical entropy

If it is true that the partition function contains all thermodynamic information, then it must be possible to use it to calculate the entropy as well as the internal energy. Because we know (from Section 3.2) that entropy is related to the dispersal of energy and that the partition function is a measure of the number of thermally accessible states, we can be confident that the two are indeed related.

We shall develop the relation between the entropy and the partition function in two stages. In *Further information* 16.2, we justify one of the most celebrated equations in statistical thermodynamics, the **Boltzmann formula** for the entropy:

$$S = k \ln W \qquad [16.34]$$

In this expression, W is the weight of the most probable configuration of the system. In the second stage, we express W in terms of the partition function.

The statistical entropy behaves in exactly the same way as the thermodynamic entropy. Thus, as the temperature is lowered, the value of W, and hence of S, decreases because fewer configurations are compatible with the total energy. In the limit $T \to 0$, $W = 1$, so $\ln W = 0$, because only one configuration (every molecule in the lowest level) is compatible with $E = 0$. It follows that $S \to 0$ as $T \to 0$, which is compatible with the Third Law of thermodynamics, that the entropies of all perfect crystals approach the same value as $T \to 0$ (Section 3.4).

Now we relate the Boltzmann formula for the entropy to the partition function. To do so, we substitute the expression for $\ln W$ given in eqn 16.3 into eqn 16.34 and, as shown in the *Justification* below, obtain

$$S = \frac{U - U(0)}{T} + Nk \ln q \qquad (16.35)$$

Justification 16.4 *The statistical entropy*

The first stage is to use eqn 16.3 ($\ln W = N \ln N - \sum_i n_i \ln n_i$) and $N = \sum_i n_i$ to write

$$S = k\sum_i (n_i \ln N - n_i \ln n_i) = -k\sum_i n_i \ln \frac{n_i}{N} = -Nk\sum_i p_i \ln p_i$$

where $p_i = n_i/N$, the fraction of molecules in state i. It follows from eqn 16.7 that

$$\ln p_i = -\beta\varepsilon_i - \ln q$$

and therefore that

$$S = -Nk(-\beta\sum_i p_i\varepsilon_i - \sum_i p_i \ln q) = k\beta\{U - U(0)\} + Nk \ln q$$

We have used the fact that the sum over the p_i is equal to 1 and that (from eqns 16.27 and 16.30)

$$N\sum_i p_i\varepsilon_i = \sum_i Np_i\varepsilon_i = \sum_i Np_i\varepsilon_i = \sum_i n_i\varepsilon_i = E = U - U(0)$$

We have already established that $\beta = 1/kT$, so eqn 16.35 immediately follows.

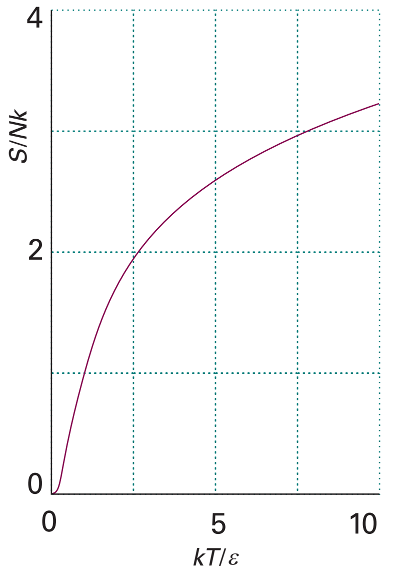

Fig. 16.12 The temperature variation of the entropy of the system shown in Fig. 16.3 (expressed here as a multiple of Nk). The entropy approaches zero as $T \rightarrow 0$, and increases without limit as $T \rightarrow \infty$.

Exploration Plot the function $\mathrm{d}S/\mathrm{d}T$, the temperature coefficient of the entropy, against kT/ε. Is there a temperature at which this coefficient passes through a maximum? If you find a maximum, explain its physical origins.

Example 16.4 *Calculating the entropy of a collection of oscillators*

Calculate the entropy of a collection of N independent harmonic oscillators, and evaluate it using vibrational data for I_2 vapour at 25°C (Example 16.3).

Method To use eqn 16.35, we use the partition function for a molecule with evenly spaced vibrational energy levels, eqn 16.12. With the partition function available, the internal energy can be found by differentiation (as in eqn 16.31a), and the two expressions then combined to give S.

Answer The molecular partition function as given in eqn 16.12 is

$$q = \frac{1}{1 - \mathrm{e}^{-\beta\varepsilon}}$$

The internal energy is obtained by using eqn 16.31a:

$$U - U(0) = -\frac{N}{q}\left(\frac{\partial q}{\partial \beta}\right)_V = \frac{N\varepsilon \mathrm{e}^{-\beta\varepsilon}}{1 - \mathrm{e}^{-\beta\varepsilon}} = \frac{N\varepsilon}{\mathrm{e}^{-\beta\varepsilon} - 1}$$

The entropy is therefore

$$S = Nk\left\{\frac{\beta\varepsilon}{\mathrm{e}^{\beta\varepsilon} - 1} - \ln(1 - \mathrm{e}^{\beta\varepsilon})\right\}$$

This function is plotted in Fig. 16.12. For I_2 at 25°C, $\beta\varepsilon = 1.036$ (Example 16.3), so $S_m = 8.38\ \mathrm{J\ K^{-1}\ mol^{-1}}$.

Self-test 16.6 Evaluate the molar entropy of N two-level systems and plot the resulting expression. What is the entropy when the two states are equally thermally accessible?

[$S/Nk = \beta\varepsilon/(1 + \mathrm{e}^{\beta\varepsilon}) + \ln(1 + \mathrm{e}^{-\beta\varepsilon})$; see Fig. 16.13; $S = Nk \ln 2$]

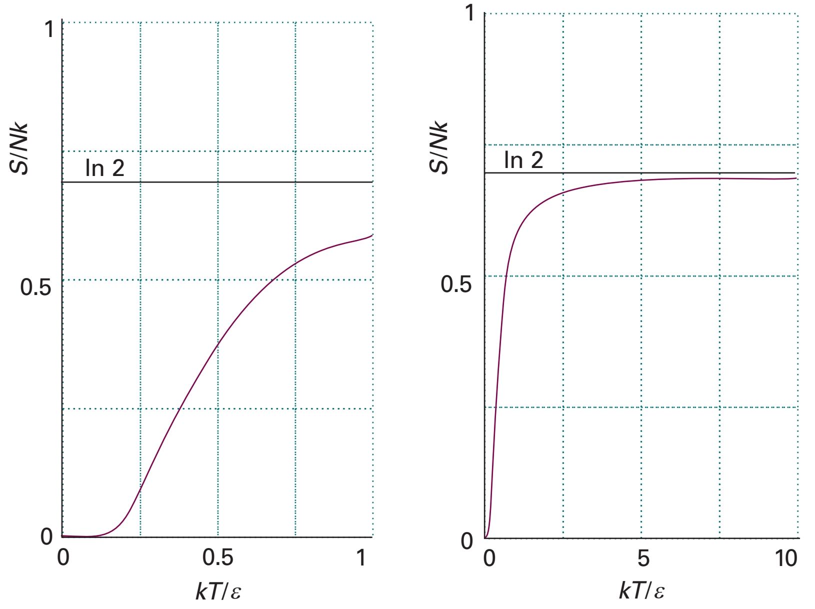

Fig. 16.13 The temperature variation of the entropy of a two-level system (expressed as a multiple of Nk). As $T \rightarrow \infty$, the two states become equally populated and S approaches $Nk \ln 2$.

Exploration Draw graphs similar to those in Fig. 16.13 for a three-level system with levels 0, ε, and 2ε.

The canonical partition function

In this section we see how to generalize our conclusions to include systems composed of interacting molecules. We shall also see how to obtain the molecular partition function from the more general form of the partition function developed here.

16.5 The canonical ensemble

The crucial new concept we need when treating systems of interacting particles is the 'ensemble'. Like so many scientific terms, the term has basically its normal meaning of 'collection', but it has been sharpened and refined into a precise significance.

(a) The concept of ensemble

To set up an ensemble, we take a closed system of specified volume, composition, and temperature, and think of it as replicated $\tilde{N}$ times (Fig. 16.14). All the identical closed systems are regarded as being in thermal contact with one another, so they can exchange energy. The total energy of all the systems is $\tilde{E}$ and, because they are in thermal equilibrium with one another, they all have the same temperature, T. This imaginary collection of replications of the actual system with a common temperature is called the **canonical ensemble**.

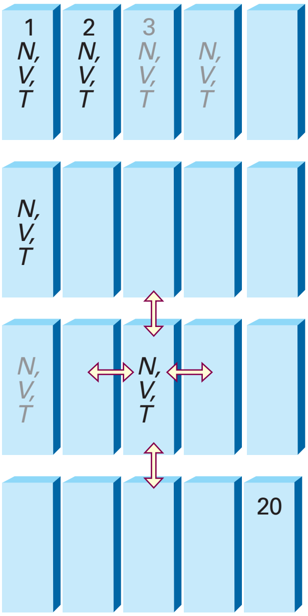

Fig. 16.14 A representation of the canonical ensemble, in this case for $\tilde{N} = 20$. The individual replications of the actual system all have the same composition and volume. They are all in mutual thermal contact, and so all have the same temperature. Energy may be transferred between them as heat, and so they do not all have the same energy. The total energy $\tilde{E}$ of all 20 replications is a constant because the ensemble is isolated overall.

The word 'canon' means 'according to a rule'. There are two other important ensembles. In the **microcanonical ensemble** the condition of constant temperature is replaced by the requirement that all the systems should have exactly the same energy: each system is individually isolated. In the **grand canonical ensemble** the volume and temperature of each system is the same, but they are open, which means that matter can be imagined as able to pass between the systems; the composition of each one may fluctuate, but now the chemical potential is the same in each system:

Microcanonical ensemble: N, V, E common

Canonical ensemble: N, V, T common

Grand canonical ensemble: μ, V, T common

The important point about an ensemble is that it is a collection of *imaginary* replications of the system, so we are free to let the number of members be as large as we like; when appropriate, we can let $\tilde{N}$ become infinite. The number of members of the ensemble in a state with energy E_i is denoted $\tilde{n}_i$, and we can speak of the configuration of the ensemble (by analogy with the configuration of the system used in Section 16.1) and its weight, $\tilde{W}$. Note that $\tilde{N}$ is unrelated to N, the number of molecules in the actual system; $\tilde{N}$ is the number of imaginary replications of that system.

(b) Dominating configurations

Just as in Section 16.1, some of the configurations of the ensemble will be very much more probable than others. For instance, it is very unlikely that the whole of the total energy, $\tilde{E}$, will accumulate in one system. By analogy with the earlier discussion, we can anticipate that there will be a dominating configuration, and that we can evaluate the thermodynamic properties by taking the average over the ensemble using that single, most probable, configuration. In the **thermodynamic limit** of $\tilde{N} \rightarrow \infty$, this dominating configuration is overwhelmingly the most probable, and it dominates the properties of the system virtually completely.

The quantitative discussion follows the argument in Section 16.1 with the modification that N and n_i are replaced by $\tilde{N}$ and $\tilde{n}_i$. The weight of a configuration $\{\tilde{n}_0, \tilde{n}_1, \ldots\}$ is

$$\tilde{W} = \frac{\tilde{N}!}{\tilde{n}_0!\tilde{n}_1!\ldots} \tag{16.36}$$

The configuration of greatest weight, subject to the constraints that the total energy of the ensemble is constant at $\tilde{E}$ and that the total number of members is fixed at $\tilde{N}$, is given by the **canonical distribution**:

$$\frac{\tilde{n}_i}{\tilde{N}} = \frac{e^{-\beta E_i}}{Q} \qquad Q = \sum_i e^{-\beta E_i} \tag{16.37}$$

The quantity Q, which is a function of the temperature, is called the **canonical partition function**.

(c) Fluctuations from the most probable distribution

The canonical distribution in eqn 16.37 is only apparently an exponentially decreasing function of the energy of the system. We must appreciate that eqn 16.37 gives the probability of occurrence of members in a single state i of the entire system of energy E_i. There may in fact be numerous states with almost identical energies. For example, in a gas the identities of the molecules moving slowly or quickly can change without necessarily affecting the total energy. The density of states, the number of states in an energy range divided by the width of the range (Fig. 16.15), is a very sharply increasing function of energy. It follows that the probability of a member of an ensemble having a specified energy (as distinct from being in a specified state) is given by eqn 16.37, a sharply decreasing function, multiplied by a sharply increasing function (Fig. 16.16). Therefore, the overall distribution is a sharply peaked function. We conclude that most members of the ensemble have an energy very close to the mean value.

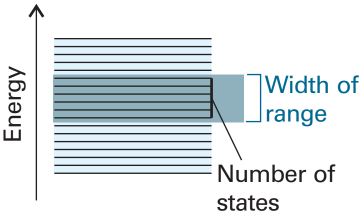

Fig. 16.15 The energy density of states is the number of states in an energy range divided by the width of the range.

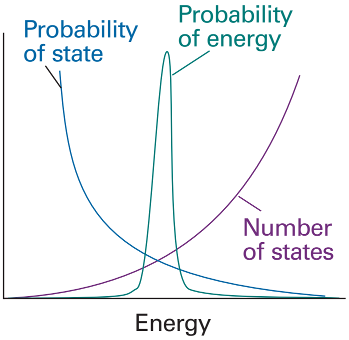

Fig. 16.16 To construct the form of the distribution of members of the canonical ensemble in terms of their energies, we multiply the probability that any one is in a state of given energy, eqn 16.39, by the number of states corresponding to that energy (a steeply rising function). The product is a sharply peaked function at the mean energy, which shows that almost all the members of the ensemble have that energy.

16.6 The thermodynamic information in the partition function

Like the molecular partition function, the canonical partition function carries all the thermodynamic information about a system. However, Q is more general than q because it does not assume that the molecules are independent. We can therefore use Q to discuss the properties of condensed phases and real gases where molecular interactions are important.

(a) The internal energy

If the total energy of the ensemble is $\tilde{E}$, and there are $\tilde{N}$ members, the average energy of a member is $E = \tilde{E}/\tilde{N}$. We use this quantity to calculate the internal energy of the system in the limit of $\tilde{N}$ (and $\tilde{E}$) approaching infinity:

$$U = U(0) + E = U(0) + \tilde{E}/\tilde{N} \qquad \text{as} \qquad \tilde{N} \to \infty \tag{16.38}$$

The fraction, $\tilde{p}_i$, of members of the ensemble in a state i with energy E_i is given by the analogue of eqn 16.7 as

$$\tilde{p}_i = \frac{e^{-\beta E_i}}{Q} \tag{16.39}$$

It follows that the internal energy is given by

$$U = U(0) + \sum_i \tilde{p}_i E_i = U(0) + \frac{1}{Q}\sum_i E_i e^{-\beta E_i} \tag{16.40}$$

By the same argument that led to eqn 16.31,

$$U = U(0) - \frac{1}{Q}\left(\frac{\partial Q}{\partial \beta}\right)_V = U(0) - \left(\frac{\partial \ln Q}{\partial \beta}\right)_V \tag{16.41}$$

(b) The entropy

The total weight, $\tilde{W}$, of a configuration of the ensemble is the product of the average weight W of each member of the ensemble, $\tilde{W} = W^{\tilde{N}}$. Hence, we can calculate S from

$$S = k \ln W = k \ln \tilde{W}^{1/\tilde{N}} = \frac{k}{\tilde{N}} \ln \tilde{W} \tag{16.42}$$

It follows, by the same argument used in Section 16.4, that

$$S = \frac{U - U(0)}{T} + k \ln Q \tag{16.43}$$

16.7 Independent molecules

We shall now see how to recover the molecular partition function from the more general canonical partition function when the molecules are independent. When the molecules are independent and distinguishable (in the sense to be described), the relation between Q and q is

$$Q = q^N \tag{16.44}$$

Justification 16.5 *The relation between Q and q*

The total energy of a collection of N independent molecules is the sum of the energies of the molecules. Therefore, we can write the total energy of a state i of the system as

$$E_i = \varepsilon_i(1) + \varepsilon_i(2) + \cdots + \varepsilon_i(N)$$

In this expression, $\varepsilon_i(1)$ is the energy of molecule 1 when the system is in the state i, $\varepsilon_i(2)$ the energy of molecule 2 when the system is in the same state i, and so on. The canonical partition function is then

$$Q = \sum_i \mathrm{e}^{-\beta\varepsilon_i(1) - \beta\varepsilon_i(2) - \cdots - \beta\varepsilon_i(N)}$$

The sum over the states of the system can be reproduced by letting each molecule enter all its own individual states (although we meet an important proviso shortly). Therefore, instead of summing over the states i of the system, we can sum over all the individual states i of molecule 1, all the states i of molecule 2, and so on. This rewriting of the original expression leads to

$$Q = \left(\sum_i \mathrm{e}^{-\beta\varepsilon_i}\right)\left(\sum_i \mathrm{e}^{-\beta\varepsilon_i}\right)\cdots\left(\sum_i \mathrm{e}^{-\beta\varepsilon_i}\right) = \left(\sum_i \mathrm{e}^{-\beta\varepsilon_i}\right)^N = q^N$$

(a) Distinguishable and indistinguishable molecules

If all the molecules are identical and free to move through space, we cannot distinguish them and the relation $Q = q^N$ is not valid. Suppose that molecule 1 is in some state a, molecule 2 is in b, and molecule 3 is in c, then one member of the ensemble has an energy $E = \varepsilon_a + \varepsilon_b + \varepsilon_c$. This member, however, is indistinguishable from one formed by putting molecule 1 in state b, molecule 2 in state c, and molecule 3 in state a, or some other permutation. There are six such permutations in all, and $N!$ in general. In the case of indistinguishable molecules, it follows that we have counted too many states in going from the sum over system states to the sum over molecular states, so writing $Q = q^N$ overestimates the value of Q. The detailed argument is quite involved, but at all except very low temperatures it turns out that the correction factor is $1/N!$. Therefore:

- For distinguishable independent molecules: $Q = q^N$ (16.45a)
- For indistinguishable independent molecules: $Q = q^N/N!$ (16.45b)

For molecules to be indistinguishable, they must be of the same kind: an Ar atom is never indistinguishable from a Ne atom. Their identity, however, is not the only criterion. Each identical molecule in a crystal lattice, for instance, can be 'named' with a set of coordinates. Identical molecules in a lattice can therefore be treated as distinguishable because their sites are distinguishable, and we use eqn 16.45a. On the other hand, identical molecules in a gas are free to move to different locations, and there is no way of keeping track of the identity of a given molecule; we therefore use eqn 16.45b.

(b) The entropy of a monatomic gas

An important application of the previous material is the derivation (as shown in the *Justification* below) of the **Sackur–Tetrode equation** for the entropy of a monatomic gas:

$$S = nR\ln\left(\frac{e^{5/2}V}{nN_A\Lambda^3}\right) \qquad \Lambda = \frac{h}{(2\pi mkT)^{1/2}} \tag{16.46a}$$

This equation implies that the molar entropy of a perfect gas of high molar mass is greater than one of low molar mass under the same conditions (because the former has more thermally accessible translational states). Because the gas is perfect, we can use the relation $V = nRT/p$ to express the entropy in terms of the pressure as

$$S = nR\ln\left(\frac{e^{5/2}kT}{p\Lambda^3}\right) \tag{16.46b}$$

Justification 16.6 *The Sackur–Tetrode equation*

For a gas of independent molecules, Q may be replaced by $q^N/N!$, with the result that eqn 16.43 becomes

$$S = \frac{U - U(0)}{T} + Nk\ln q - k\ln N!$$

Because the number of molecules ($N = nN_A$) in a typical sample is large, we can use Stirling's approximation (eqn 16.2) to write

$$S = \frac{U - U(0)}{T} + nR\ln q - nR\ln N + nR$$

The only mode of motion for a gas of atoms is translation, and the partition function is $q = V/\Lambda^3$ (eqn 16.19), where Λ is the thermal wavelength. The internal energy is given by eqn 16.32, so the entropy is

$$S = \tfrac{3}{2}nR + nR\left(\ln\frac{V}{\Lambda^3} - \ln nN_A + 1\right) = nR\left(\ln e^{3/2} + \ln\frac{V}{\Lambda^3} - \ln nN_A + \ln e\right)$$

which rearranges into eqn 16.46.

Example 16.5 *Using the Sackur–Tetrode equation*

Calculate the standard molar entropy of gaseous argon at 25°C.

Method To calculate the molar entropy, S_m, from eqn 16.46b, divide both sides by n. To calculate the standard molar entropy, $S_m^\ominus$, set $p = p^\ominus$ in the expression for S_m:

$$S_m^\ominus = R \ln\left(\frac{e^{5/2}kT}{p^\ominus \Lambda^3}\right)$$

Answer The mass of an Ar atom is $m = 39.95$ u. At 25°C, its thermal wavelength is 16.0 pm (by the same kind of calculation as in *Illustration* 16.3). Therefore,

$$S_m^\ominus = R \ln\left\{\frac{e^{5/2} \times (4.12 \times 10^{-21}\ \mathrm{J})}{(10^5\ \mathrm{N\ m^{-2}}) \times (1.60 \times 10^{-11}\ \mathrm{m})^3}\right\} = 18.6R = 155\ \mathrm{J\ K^{-1}\ mol^{-1}}$$

We can anticipate, on the basis of the number of accessible states for a lighter molecule, that the standard molar entropy of Ne is likely to be smaller than for Ar; its actual value is 17.60R at 298 K.

Self-test 16.7 Calculate the translational contribution to the standard molar entropy of H_2 at 25°C. [14.2R]

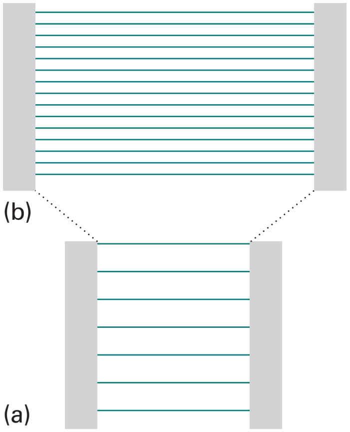

Fig. 16.17 As the width of a container is increased (going from (a) to (b)), the energy levels become closer together (as $1/L^2$), and as a result more are thermally accessible at a given temperature. Consequently, the entropy of the system rises as the container expands.

The Sackur–Tetrode equation implies that, when a monatomic perfect gas expands isothermally from V_i to V_f, its entropy changes by

$$\Delta S = nR \ln(aV_f) - nR \ln(aV_i) = nR \ln \frac{V_f}{V_i} \qquad (16.47)$$

where aV is the collection of quantities inside the logarithm of eqn 16.46a. This is exactly the expression we obtained by using classical thermodynamics (Example 3.1). Now, though, we see that that classical expression is in fact a consequence of the increase in the number of accessible translational states when the volume of the container is increased (Fig. 16.17).

Checklist of key ideas

☐ 1. The instantaneous configuration of a system of N molecules is the specification of the set of populations $n_0, n_1, \ldots$ of the energy levels $\varepsilon_0, \varepsilon_1, \ldots$. The weight W of a configuration is given by $W = N!/n_0!n_1! \ldots$.

☐ 2. The Boltzmann distribution gives the numbers of molecules in each state of a system at any temperature: $N_i = Ne^{-\beta\varepsilon_i}/q$, $\beta = 1/kT$.

☐ 3. The partition function is defined as $q = \sum_j e^{-\beta\varepsilon_j}$ and is an indication of the number of thermally accessible states at the temperature of interest.

☐ 4. The internal energy is $U(T) = U(0) + E$, with $E = -(N/q)(\partial q/\partial\beta)_V = -N(\partial \ln q/\partial\beta)_V$.

☐ 5. The Boltzmann formula for the entropy is $S = k \ln W$, where W is the number of different ways in which the molecules of a system can be arranged while keeping the same total energy.

☐ 6. The entropy in terms of the partition function is $S = \{U - U(0)\}/T + Nk \ln q$ (distinguishable molecules) or $S = \{U - U(0)\}/T + Nk \ln q - Nk(\ln N - 1)$ (indistinguishable molecules).

☐ 7. The canonical ensemble is an imaginary collection of replications of the actual system with a common temperature.

☐ 8. The canonical distribution is given by $\tilde{n}_i/\tilde{N} = e^{-\beta E_i}/\sum_j e^{-\beta E_j}$. The canonical partition function, $Q = \sum_i e^{-\beta E_i}$.

☐ 9. The internal energy and entropy of an ensemble are, respectively, $U = U(0) - (\partial \ln Q/\partial\beta)_V$ and $S = \{U - U(0)\}/T + k \ln Q$.

☐ 10. For distinguishable independent molecules we write $Q = q^N$. For indistinguishable independent molecules we write $Q = q^N/N!$.

☐ 11. The Sackur–Tetrode equation, eqn 16.46, is an expression for the entropy of a monatomic gas.

Further reading

Articles and texts

D. Chandler, *Introduction to modern statistical mechanics*. Oxford University Press (1987).

D.A. McQuarrie and J.D. Simon, *Molecular thermodynamics*. University Science Books, Sausalito (1999).

K.E. van Holde, W.C. Johnson, and P.S. Ho, *Principles of physical biochemistry*. Prentice Hall, Upper Saddle River (1998).

J. Wisniak, Negative absolute temperatures, a novelty. *J. Chem. Educ.* **77**, 518 (2000).

Further information

Further information 16.1 *The Boltzmann distribution*

We remarked in Section 16.1 that ln W is easier to handle than W. Therefore, to find the form of the Boltzmann distribution, we look for the condition for ln W being a maximum rather than dealing directly with W. Because ln W depends on all the n_i, when a configuration changes and the n_i change to $n_i + \mathrm{d}n_i$, the function ln W changes to ln W + d ln W, where

$$\mathrm{d}\ln W = \sum_i \left(\frac{\partial \ln W}{\partial n_i}\right)\mathrm{d}n_i$$

All this expression states is that a change in ln W is the sum of contributions arising from changes in each value of n_i. At a maximum, d ln W = 0. However, when the n_i change, they do so subject to the two constraints

$$\sum_i \varepsilon_i \mathrm{d}n_i = 0 \qquad \sum_i \mathrm{d}n_i = 0 \tag{16.48}$$

The first constraint recognizes that the total energy must not change, and the second recognizes that the total number of molecules must not change. These two constraints prevent us from solving d ln W = 0 simply by setting all $(\partial \ln W/\partial n_i) = 0$ because the dn_i are not all independent.

The way to take constraints into account was devised by the French mathematician Lagrange, and is called the **method of undetermined multipliers**. The technique is described in *Appendix* 2. All we need here is the rule that a constraint should be multiplied by a constant and then added to the main variation equation. The variables are then treated as though they were all independent, and the constants are evaluated at the end of the calculation.

We employ the technique as follows. The two constraints in eqn 16.48 are multiplied by the constants $-\beta$ and α, respectively (the minus sign in $-\beta$ has been included for future convenience), and then added to the expression for d ln W:

$$\begin{aligned}\mathrm{d}\ln W &= \sum_i \left(\frac{\partial \ln W}{\partial n_i}\right)\mathrm{d}n_i + \alpha\sum_i \mathrm{d}n_i - \beta\sum_i \varepsilon_i \mathrm{d}n_i \\ &= \sum_i \left\{\left(\frac{\partial \ln W}{\partial n_i}\right) + \alpha - \beta\varepsilon_i\right\}\mathrm{d}n_i\end{aligned}$$

All the dn_i are now treated as independent. Hence the only way of satisfying d ln W = 0 is to require that, for each i,

$$\frac{\partial \ln W}{\partial n_i} + \alpha - \beta\varepsilon_i = 0 \tag{16.49}$$

when the n_i have their most probable values.

Differentiation of ln W as given in eqn 16.3 with respect to n_i gives

$$\frac{\partial \ln W}{\partial n_i} = \frac{\partial(N\ln N)}{\partial n_i} - \sum_j \frac{\partial(n_j \ln n_j)}{\partial n_i}$$

The derivative of the first term is obtained as follows:

$$\begin{aligned}\frac{\partial(N\ln N)}{\partial n_i} &= \left(\frac{\partial N}{\partial n_i}\right)\ln N + N\left(\frac{\partial \ln N}{\partial n_i}\right) \\ &= \ln N + \frac{\partial N}{\partial n_i} = \ln N + 1\end{aligned}$$

The ln N in the first term on the right in the second line arises because $N = n_1 + n_2 + \cdots$ and so the derivative of N with respect to any of the n_i is 1: that is, $\partial N/\partial n_i = 1$. The second term on the right in the second line arises because $\partial(\ln N)/\partial n_i = (1/N)\partial N/\partial n_i$. The final 1 is then obtained in the same way as in the preceding remark, by using $\partial N/\partial n_i = 1$.

For the derivative of the second term we first note that

$$\frac{\partial \ln n_j}{\partial n_i} = \frac{1}{n_j}\left(\frac{\partial n_j}{\partial n_i}\right)$$

Morever, if $i \neq j$, n_j is independent of n_i, so $\partial n_j/\partial n_i = 0$. However, if $i = j$,

$$\frac{\partial n_j}{\partial n_i} = \frac{\partial n_j}{\partial n_j} = 1$$

Therefore,

$$\frac{\partial n_j}{\partial n_i} = \delta_{ij}$$

with δ_{ij} the Kronecker delta ($\delta_{ij} = 1$ if $i = j$, $\delta_{ij} = 0$ otherwise). Then

$$\begin{aligned}\sum_j \frac{\partial(n_j \ln n_j)}{\partial n_i} &= \sum_j \left\{ \left(\frac{\partial n_j}{\partial n_i}\right) \ln n_j + n_j \left(\frac{\partial \ln n_j}{\partial n_i}\right) \right\} \\ &= \sum_j \left\{ \left(\frac{\partial n_j}{\partial n_i}\right) \ln n_j + \left(\frac{\partial n_j}{\partial n_i}\right) \right\} \\ &= \sum_j \left(\frac{\partial n_j}{\partial n_i}\right)(\ln n_j + 1) \\ &= \sum_j \delta_{ij}(\ln n_j + 1) = \ln n_i + 1\end{aligned}$$

and therefore

$$\frac{\partial \ln W}{\partial n_i} = -(\ln n_i + 1) + (\ln N + 1) = -\ln \frac{n_i}{N}$$

It follows from eqn 16.49 that

$$-\ln \frac{n_i}{N} + \alpha - \beta\varepsilon_i = 0$$

and therefore that

$$\frac{n_i}{N} = e^{\alpha - \beta\varepsilon_i}$$

At this stage we note that

$$N = \sum_i n_i = \sum_i N e^{\alpha - \beta\varepsilon_i} = N e^{\alpha} \sum_i e^{\beta\varepsilon_i}$$

Because the N cancels on each side of this equality, it follows that

$$e^{\alpha} = \frac{1}{\sum_j e^{-\beta\varepsilon_j}} \qquad (16.50)$$

and

$$\frac{n_i}{N} = e^{\alpha - \beta\varepsilon_i} = e^{\alpha} e^{-\beta\varepsilon_i} = \frac{1}{\sum_j e^{-\beta\varepsilon_j}} e^{-\beta\varepsilon_i}$$

which is eqn 16.6a.

Further information 16.2 *The Boltzmann formula*

A change in the internal energy

$$U = U(0) + \sum_i n_i \varepsilon_i \qquad (16.51)$$

may arise from either a modification of the energy levels of a system (when ε_i changes to $\varepsilon_i + d\varepsilon_i$) or from a modification of the populations (when n_i changes to $n_i + dn_i$). The most general change is therefore

$$dU = dU(0) + \sum_i n_i d\varepsilon_i + \sum_i \varepsilon_i dn_i \qquad (16.52)$$

Because the energy levels do not change when a system is heated at constant volume (Fig. 16.18), in the absence of all changes other than heating

$$dU = \sum_i \varepsilon_i dn_i$$

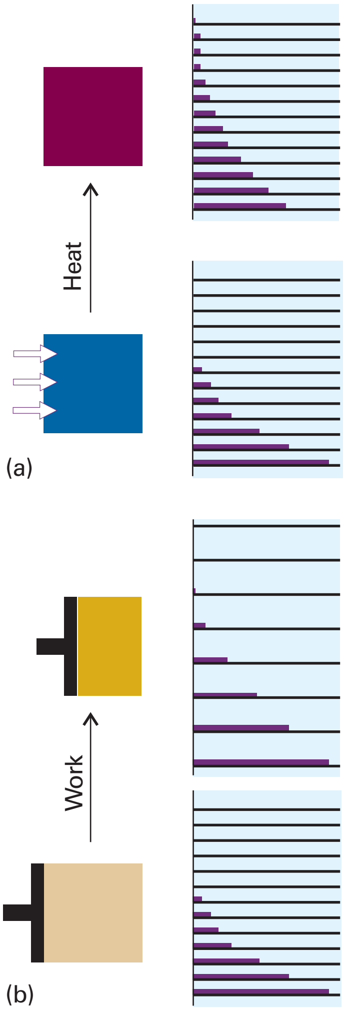

Fig. 16.18 (a) When a system is heated, the energy levels are unchanged but their populations are changed. (b) When work is done on a system, the energy levels themselves are changed. The levels in this case are the one-dimensional particle-in-a-box energy levels of Chapter 9: they depend on the size of the container and move apart as its length is decreased.

We know from thermodynamics (and specifically from eqn 3.43) that under the same conditions

$$dU = dq_{\rm rev} = TdS$$

Therefore,

$$dS = \frac{dU}{T} = k\beta \sum_i \varepsilon_i dn_i \qquad (16.53)$$

For changes in the most probable configuration (the only one we need consider), we rearrange eqn 16.49 to

$$\beta\varepsilon_i = \frac{\partial \ln W}{\partial n_i} + \alpha$$

and find that

$$\mathrm{d}S = k\sum_i \left(\frac{\partial \ln W}{\partial n_i}\right)\mathrm{d}n_i + k\alpha\sum_i \mathrm{d}n_i$$

But because the number of molecules is constant, the sum over the $\mathrm{d}n_i$ is zero. Hence

$$\mathrm{d}S = k\sum_i \left(\frac{\partial \ln W}{\partial n_i}\right)\mathrm{d}n_i = k(\mathrm{d}\ln W)$$

This relation strongly suggests the definition $S = k\ln W$, as in eqn 16.34.

Further information 16.3 *Temperatures below zero*

The Boltzmann distribution tells us that the ratio of populations in a two-level system at a temperature T is

$$\frac{N_+}{N_-} = \mathrm{e}^{-\varepsilon/kT} \qquad (16.54)$$

where ε is the separation of the upper state N_+ and the lower state N_-. It follows that, if we can contrive the population of the upper state to exceed that of the lower state, then the temperature must have a negative value. Indeed, for a general population,

$$T = \frac{\varepsilon/k}{\ln(N_-/N_+)} \qquad (16.55)$$

and the temperature is formally negative for all $N_+ > N_-$.

All the statistical thermodynamic expressions we have derived apply to $T < 0$ as well as to $T > 0$, the difference being that states with $T < 0$ are not in thermal equilibrium and therefore have to be achieved by techniques that do not rely on the equalization of temperatures of the system and its surroundings. The Third Law of thermodynamics prohibits the achievement of absolute zero in a finite number of steps. However, it is possible to circumvent this restriction in systems that have a finite number of levels or in systems that are effectively finite because they have such weak coupling to their surroundings. The practical realization of such a system is a group of spin-$\frac{1}{2}$ nuclei that have very long relaxation times, such as the ^{19}F nuclei in cold solid LiF. Pulse techniques in NMR can achieve non-equilibrium populations (Section 15.8) as can pumping procedures in laser technologies (Section 14.5). From now on, we shall suppose that these non-equilibrium distributions have been achieved, and will concentrate on the consequences.

The expressions for q, U, and S that we have derived in this chapter are applicable to $T < 0$ as well as to $T > 0$, and are shown in Fig. 16.19. We see that q and U show sharp discontinuities on passing through zero, and $T = +0$ (corresponding to all population in the lower state) is quite distinct from $T = -0$, where all the population is in the upper state. The entropy S is continuous at $T = 0$. But all these functions are continuous if we use $\beta = 1/kT$ as the dependent variable (Fig. 16.20), which shows that β is a more natural, if less familiar, variable than T. Note that $U \rightarrow 0$ as $\beta \rightarrow \infty$ (that is, as $T \rightarrow 0$, when only the lower state is occupied) and $U \rightarrow N\varepsilon$ as $\beta \rightarrow -\infty$ (that is, as $T \rightarrow -0$); we see that a state with $T = -0$ is 'hotter' than one with $T = +0$. The entropy of the system is zero on either side of $T = 0$, and rises to $Nk\ln 2$ as $T \rightarrow \pm\infty$. At $T = +0$ only one state is accessible (the lower state), only the upper state is accessible, so the entropy is zero in each case.

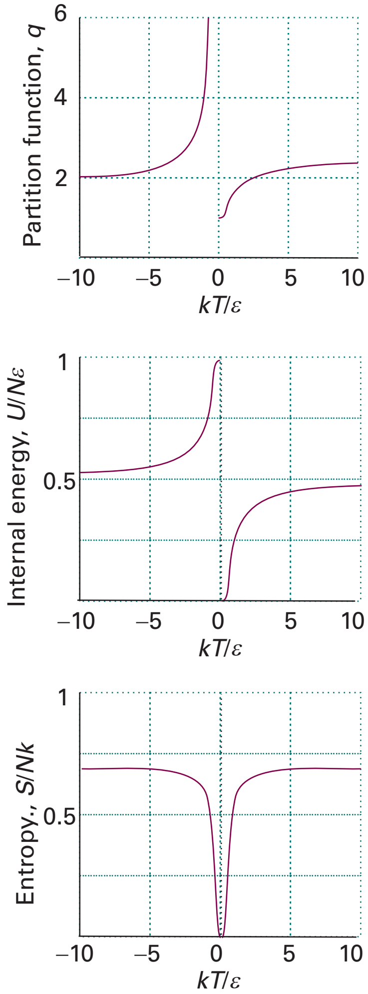

Fig. 16.19 The partition function, internal energy, and entropy of a two-level system extended to negative temperatures.

We get more insight into the dependence of thermodynamic properties on temperature by noting the thermodynamic result (Section 3.8) that $T = (\partial S/\partial U)_T$. When S is plotted against U for a two-level system (Fig. 6.21), we see that the entropy rises as energy is supplied to the system (as we would expect) provided that $T > 0$ (the thermal equilibrium regime). However, the entropy decreases as energy is supplied when $T < 0$. This conclusion is consistent with the thermodynamic definition of entropy, $\mathrm{d}S = \mathrm{d}q_{\mathrm{rev}}/T$ (where, of course, q denotes heat and not the partition function). Physically, the increase in entropy for $T > 0$ corresponds to the increasing accessibility of the upper state, and the decrease for $T < 0$ corresponds

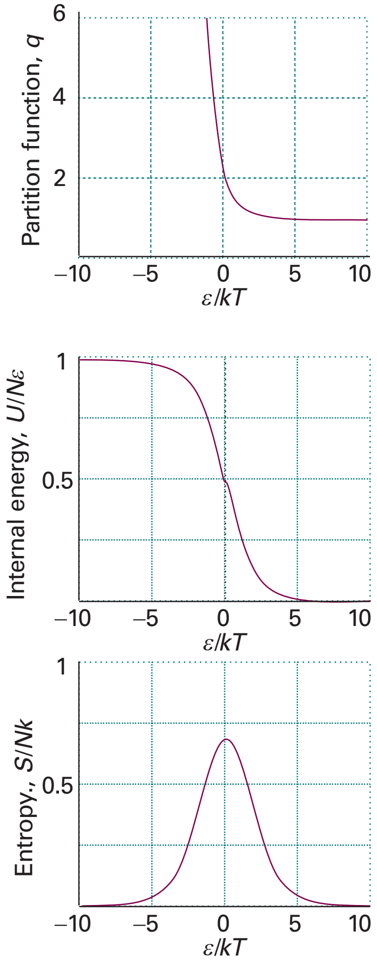

Fig. 16.20 The partition function, internal energy, and entropy of a two-level system extended to negative temperatures but plotted against $\beta = 1/kT$ (modified here to the dimensionless quantity ε/kT).

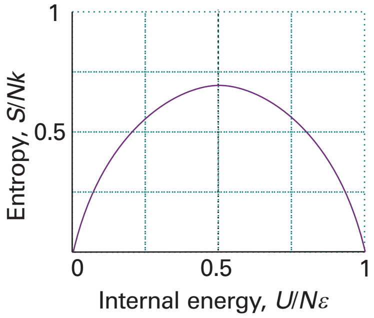

Fig. 16.21 The variation of the entropy with internal energy for a two-level system extended to negative temperatures.

to the shift towards population of the upper state alone as more energy is packed into the system.

The phenomenological laws of thermodynamics survive largely intact at negative temperatures. The First Law (in essence, the conservation of energy) is robust, and independent of how populations are distributed over states. The Second Law survives because the definition of entropy survives (as we have seen above). The efficiency of heat engines (Section 3.2), which is a direct consequence of the Second Law, is still given by $1 - T_{\text{cold}}/T_{\text{hot}}$. However, if the temperature of the cold reservoir is negative, then the efficiency of the engine may be greater than 1. This condition corresponds to the amplification of signals achieved in lasers. Alternatively, an efficiency greater than 1 implies that heat *can* be converted completely into work provided the heat is withdrawn from a reservoir at $T < 0$. If both reservoirs are at negative temperatures, then the efficiency is less than 1, as in the thermal equilibrium case treated in Chapter 3. The Third Law requires a slight amendment on account of the discontinuity of the populations across $T = 0$: it is impossible in a finite number of steps to cool any system down to +0 or to heat any system above −0.

Discussion questions

16.1 Describe the physical significance of the partition function.

16.2 Explain how the internal energy and entropy of a system composed of two levels vary with temperature.

16.3 Enumerate the ways by which the parameter β may be identified with $1/kT$.

16.4 Distinguish between the zipper and Zimm–Bragg models of the helix–coil transition.

16.5 Explain what is meant by an *ensemble* and why it is useful in statistical thermodynamics.

16.6 Under what circumstances may identical particles be regarded as distinguishable?

Exercises

16.1a What are the relative populations of the states of a two-level system when the temperature is infinite?

16.1b What is the temperature of a two-level system of energy separation equivalent to 300 cm^{-1} when the population of the upper state is one-half that of the lower state?

16.2a Calculate the translational partition function at (a) 300 K and (b) 600 K of a molecule of molar mass 120 g mol^{-1} in a container of volume 2.00 cm^3.

16.2b Calculate (a) the thermal wavelength, (b) the translational partition function of an Ar atom in a cubic box of side 1.00 cm at (i) 300 K and (ii) 3000 K.

16.3a Calculate the ratio of the translational partition functions of D_2 and H_2 at the same temperature and volume.

16.3b Calculate the ratio of the translational partition functions of xenon and helium at the same temperature and volume.

16.4a A certain atom has a threefold degenerate ground level, a non-degenerate electronically excited level at 3500 cm^{-1}, and a threefold degenerate level at 4700 cm^{-1}. Calculate the partition function of these electronic states at 1900 K.

16.4b A certain atom has a doubly degenerate ground level, a triply degenerate electronically excited level at 1250 cm^{-1}, and a doubly degenerate level at 1300 cm^{-1}. Calculate the partition function of these electronic states at 2000 K.

16.5a Calculate the electronic contribution to the molar internal energy at 1900 K for a sample composed of the atoms specified in Exercise 16.4a.

16.5b Calculate the electronic contribution to the molar internal energy at 2000 K for a sample composed of the atoms specified in Exercise 16.4b.

16.6a A certain molecule has a non-degenerate excited state lying at 540 cm^{-1} above the non-degenerate ground state. At what temperature will 10 per cent of the molecules be in the upper state?

16.6b A certain molecule has a doubly degenerate excited state lying at 360 cm^{-1} above the non-degenerate ground state. At what temperature will 15 per cent of the molecules be in the upper state?

16.7a An electron spin can adopt either of two orientations in a magnetic field, and its energies are $\pm\mu_B\mathcal{B}$, where μ_B is the Bohr magneton. Deduce an expression for the partition function and mean energy of the electron and sketch the variation of the functions with $\mathcal{B}$. Calculate the relative populations of the spin states at (a) 4.0 K, (b) 298 K when $\mathcal{B} = 1.0$ T.

16.7b A nitrogen nucleus spin can adopt any of three orientations in a magnetic field, and its energies are $0, \pm\gamma_N\hbar\mathcal{B}$, where γ_N is the magnetogyric ratio of the nucleus. Deduce an expression for the partition function and mean energy of the nucleus and sketch the variation of the functions with $\mathcal{B}$. Calculate the relative populations of the spin states at (a) 1.0 K, (b) 298 K when $\mathcal{B} = 20.0$ T.

16.8a Consider a system of distinguishable particles having only two non-degenerate energy levels separated by an energy that is equal to the value of kT at 10 K. Calculate (a) the ratio of populations in the two states at (1) 1.0 K, (2) 10 K, and (3) 100 K, (b) the molecular partition function at 10 K, (c) the molar energy at 10 K, (d) the molar heat capacity at 10 K, (e) the molar entropy at 10 K.

16.8b Consider a system of distinguishable particles having only three non-degenerate energy levels separated by an energy which is equal to the value of kT at 25.0 K. Calculate (a) the ratio of populations in the states at (1) 1.00 K, (2) 25.0 K, and (3) 100 K, (b) the molecular partition function at 25.0 K, (c) the molar energy at 25.0 K, (d) the molar heat capacity at 25.0 K, (e) the molar entropy at 25.0 K.

16.9a At what temperature would the population of the first excited vibrational state of HCl be 1/e times its population of the ground state?

16.9b At what temperature would the population of the first excited rotational level of HCl be 1/e times its population of the ground state?

16.10a Calculate the standard molar entropy of neon gas at (a) 200 K, (b) 298.15 K.

16.10b Calculate the standard molar entropy of xenon gas at (a) 100 K, (b) 298.15 K.

16.11a Calculate the vibrational contribution to the entropy of Cl_2 at 500 K given that the wavenumber of the vibration is 560 cm^{-1}.

16.11b Calculate the vibrational contribution to the entropy of Br_2 at 600 K given that the wavenumber of the vibration is 321 cm^{-1}.

16.12a Identify the systems for which it is essential to include a factor of $1/N!$ on going from Q to q: (a) a sample of helium gas, (b) a sample of carbon monoxide gas, (c) a solid sample of carbon monoxide, (d) water vapour.

16.12b Identify the systems for which it is essential to include a factor of $1/N!$ on going from Q to q: (a) a sample of carbon dioxide gas, (b) a sample of graphite, (c) a sample of diamond, (d) ice.

Problems*

Numerical problems

16.1‡ Consider a system A consisting of subsystems A_1 and A_2, for which $W_1 = 1 \times 10^{20}$ and $W_2 = 2 \times 10^{20}$. What is the number of configurations available to the combined system? Also, compute the entropies S, S_1, and S_2. What is the significance of this result?

16.2‡ Consider 1.00×10^{22} ^{4}He atoms in a box of dimensions 1.0 cm × 1.0 cm × 1.0 cm. Calculate the occupancy of the first excited level at 1.0 mK, 2.0 K, and 4.0 K. Do the same for ^{3}He. What conclusions might you draw from the results of your calculations?

---

* Problems denoted with the symbol ‡ were supplied by Charles Trapp and Carmen Giunta.

16.3‡ By what factor does the number of available configurations increase when 100 J of energy is added to a system containing 1.00 mol of particles at constant volume at 298 K?

16.4‡ By what factor does the number of available configurations increase when 20 m^3 of air at 1.00 atm and 300 K is allowed to expand by 0.0010 per cent at constant temperature?

16.5 Explore the conditions under which the 'integral' approximation for the translational partition function is not valid by considering the translational partition function of an Ar atom in a cubic box of side 1.00 cm. Estimate the temperature at which, according to the integral approximation, $q = 10$ and evaluate the exact partition function at that temperature.

16.6 A certain atom has a doubly degenerate ground level pair and an upper level of four degenerate states at 450 cm^{-1} above the ground level. In an atomic beam study of the atoms it was observed that 30 per cent of the atoms were in the upper level, and the translational temperature of the beam was 300 K. Are the electronic states of the atoms in thermal equilibrium with the translational states?

16.7 (a) Calculate the electronic partition function of a tellurium atom at (i) 298 K, (ii) 5000 K by direct summation using the following data:

Term	Degeneracy	Wavenumber/cm^{-1}
Ground	5	0
1	1	4 707
2	3	4 751
3	5	10 559

(b) What proportion of the Te atoms are in the ground term and in the term labelled 2 at the two temperatures? (c) Calculate the electronic contribution to the standard molar entropy of gaseous Te atoms.

16.8 The four lowest electronic levels of a Ti atom are: $^3F_2$, $^3F_3$, $^3F_4$, and $^5F_1$, at 0, 170, 387, and 6557 cm^{-1}, respectively. There are many other electronic states at higher energies. The boiling point of titanium is 3287°C. What are the relative populations of these levels at the boiling point? *Hint.* The degeneracies of the levels are $2J + 1$.

16.9 The NO molecule has a doubly degenerate excited electronic level 121.1 cm^{-1} above the doubly degenerate electronic ground term. Calculate and plot the electronic partition function of NO from $T = 0$ to 1000 K. Evaluate (a) the term populations and (b) the electronic contribution to the molar internal energy at 300 K. Calculate the electronic contribution to the molar entropy of the NO molecule at 300 K and 500 K.

16.10‡ J. Sugar and A. Musgrove (*J. Phys. Chem. Ref. Data* **22**, 1213 (1993)) have published tables of energy levels for germanium atoms and cations from Ge^+ to Ge^{+31}. The lowest-lying energy levels in neutral Ge are as follows:

	$^3P_0$	$^3P_1$	$^3P_2$	$^1D_2$	$^1S_0$
E/cm^{-1}	0	557.1	1410.0	7125.3	16 367.3

Calculate the electronic partition function at 298 K and 1000 K by direct summation. *Hint.* The degeneracy of a level is $2J + 1$.

16.11 Calculate, by explicit summation, the vibrational partition function and the vibrational contribution to the molar internal energy of I_2 molecules at (a) 100 K, (b) 298 K given that its vibrational energy levels lie at the following wavenumbers above the zero-point energy level: 0, 213.30, 425.39, 636.27, 845.93 cm^{-1}. What proportion of I_2 molecules are in the ground and first two excited levels at the two temperatures? Calculate the vibrational contribution to the molar entropy of I_2 at the two temperatures.

16.12‡ (a) The standard molar entropy of graphite at 298, 410, and 498 K is 5.69, 9.03, and 11.63 J K^{-1} mol^{-1}, respectively. If 1.00 mol C(graphite) at 298 K is surrounded by thermal insulation and placed next to 1.00 mol C(graphite) at 498 K, also insulated, how many configurations are there altogether for the combined but independent systems? (b) If the same two samples are now placed in thermal contact and brought to thermal equilibrium, the final temperature will be 410 K. (Why might the final temperature not be the average?) How many configurations are there now in the combined system? Neglect any volume changes. (c) Demonstrate that this process is spontaneous.

Theoretical problems

16.13 A sample consisting of five molecules has a total energy 5ε. Each molecule is able to occupy states of energy $j\varepsilon$, with $j = 0, 1, 2, \ldots$. (a) Calculate the weight of the configuration in which the molecules are distributed evenly over the available states. (b) Draw up a table with columns headed by the energy of the states and write beneath them all configurations that are consistent with the total energy. Calculate the weights of each configuration and identify the most probable configurations.

16.14 A sample of nine molecules is numerically tractable but on the verge of being thermodynamically significant. Draw up a table of configurations for $N = 9$, total energy 9ε in a system with energy levels $j\varepsilon$ (as in Problem 16.13). Before evaluating the weights of the configurations, guess (by looking for the most 'exponential' distribution of populations) which of the configurations will turn out to be the most probable. Go on to calculate the weights and identify the most probable configuration.

16.15 The most probable configuration is characterized by a parameter we know as the 'temperature'. The temperatures of the system specified in Problems 16.13 and 16.14 must be such as to give a mean value of ε for the energy of each molecule and a total energy $N\varepsilon$ for the system. (a) Show that the temperature can be obtained by plotting p_j against j, where p_j is the (most probable) fraction of molecules in the state with energy $j\varepsilon$. Apply the procedure to the system in Problem 16.14. What is the temperature of the system when ε corresponds to 50 cm^{-1}? (b) Choose configurations other than the most probable, and show that the same procedure gives a worse straight line, indicating that a temperature is not well-defined for them.

16.16 A certain molecule can exist in either a non-degenerate singlet state or a triplet state (with degeneracy 3). The energy of the triplet exceeds that of the singlet by ε. Assuming that the molecules are distinguishable (localized) and independent, (a) obtain the expression for the molecular partition function. (b) Find expressions in terms of ε for the molar energy, molar heat capacity, and molar entropy of such molecules and calculate their values at $T = \varepsilon/k$.

16.17 Consider a system with energy levels $\varepsilon_j = j\varepsilon$ and N molecules. (a) Show that if the mean energy per molecule is $a\varepsilon$, then the temperature is given by

$$\beta = \frac{1}{\varepsilon}\ln\left(1 + \frac{1}{a}\right)$$

Evaluate the temperature for a system in which the mean energy is ε, taking ε equivalent to 50 cm^{-1}. (b) Calculate the molecular partition function q for the system when its mean energy is $a\varepsilon$. (c) Show that the entropy of the system is

$$S/k = (1 + a)\ln(1 + a) - a\ln a$$

and evaluate this expression for a mean energy ε.

16.18 Consider Stirling's approximation for ln $N!$ in the derivation of the Boltzmann distribution. What difference would it make if (a) a cruder approximation, $N! = N^N$, (b) the better approximation in *Comment* 16.2 were used instead?

16.19‡ For gases, the canonical partition function, Q, is related to the molecular partition function q by $Q = q^N/N!$. Use the expression for q and general thermodynamic relations to derive the perfect gas law $pV = nRT$.

Applications: to atmospheric science, astrophysics, and biochemistry

16.20‡ Obtain the barometric formula (Problem 1.27) from the Boltzmann distribution. Recall that the potential energy of a particle at height h above the surface of the Earth is mgh. Convert the barometric formula from pressure to number density, $\mathcal{N}$. Compare the relative number densities, $\mathcal{N}(h)/\mathcal{N}(0)$, for O_2 and H_2O at $h = 8.0$ km, a typical cruising altitude for commercial aircraft.

16.21‡ Planets lose their atmospheres over time unless they are replenished. A complete analysis of the overall process is very complicated and depends upon the radius of the planet, temperature, atmospheric composition, and other factors. Prove that the atmosphere of planets cannot be in an equilibrium state by demonstrating that the Boltzmann distribution leads to a uniform finite number density as $r \to \infty$. *Hint.* Recall that in a gravitational field the potential energy is $V(r) = -GMm/r$, where G is the gravitational constant, M is the mass of the planet, and m the mass of the particle.

16.22‡ Consider the electronic partition function of a perfect atomic hydrogen gas at a density of 1.99×10^{-4} kg m^{-3} and 5780 K. These are the mean conditions within the Sun's photosphere, the surface layer of the Sun that is about 190 km thick. (a) Show that this partition function, which involves a sum over an infinite number of quantum states that are solutions to the Schrödinger equation for an isolated atomic hydrogen atom, is infinite. (b) Develop a theoretical argument for truncating the sum and estimate the maximum number of quantum states that contribute to the sum. (c) Calculate the equilibrium probability that an atomic hydrogen electron is in each quantum state. Are there any general implications concerning electronic states that will be observed for other atoms and molecules? Is it wise to apply these calculations in the study of the Sun's photosphere?

16.23 Consider a protein P with four distinct sites, with each site capable of binding one ligand L. Show that the possible varieties (configurations) of the species PL_i (with PL_0 denoting P) are given by the binomial coefficients $C(4,i)$.

16.24 Complete some of the derivations in the discussion of the helix–coil transition in polypeptides (*Impact* I16.1). (a) Show that, within the tenets of the zipper model,

$$q = 1 + \sum_{i=1}^{n} Z(n,i)\sigma s^i$$

and that $Z(n,i) = n - i + 1$ is the number of ways in which an allowed state with a number i of c amino acids can be formed. (b) Using the zipper model, show that $\theta = (1/n)\mathrm{d}(\ln q)/\mathrm{d}(\ln s)$. *Hint.* As a first step, show that $\sum_i i(n-i+1)\sigma s^i = s(\mathrm{d}q/\mathrm{d}s)$.

16.25 Here you will use the zipper model discussed in *Impact* I16.1 to explore the helix–coil transition in polypeptides.(a) Investigate the effect of the parameter s on the distribution of random coil segments in a polypeptide with $n = 20$ by plotting p_i, the fraction of molecules with a number i of amino acids in a coil region, against i for $s = 0.8$, 1.0, and 1.5, with $\sigma = 5.0 \times 10^{-2}$. Discuss the significance of any effects you discover. (b) The average value of i given by $\langle i \rangle = \sum_i ip_i$. Use the results of the zipper model to calculate $\langle i \rangle$ for all the combinations of s and σ used in Fig. 16.10 and part (a).