16

THE CANONICAL FORMALISM: STATISTICAL MECHANICS IN HELMHOLTZ REPRESENTATION

16-1 THE PROBABILITY DISTRIBUTION

The microcanonical formalism of the preceding chapter is simple in principle, but it is computationally feasible only for a few highly idealized models. The combinatorial calculation of the number of ways that a given amount of energy can be distributed in arbitrarily sized "boxes" is generally beyond our mathematical capabilities. The solution is to remove the limitation on the amount of energy available—to consider a system in contact with a thermal reservoir rather than an isolated system. The statistical mechanics of a system in contact with a thermal reservoir may be viewed as statistical mechanics "in Helmholtz representation"; or, in the parlance of the field, "in canonical formalism."

States of all energies, from zero to arbitrarily large energies, are available to a system in contact with a thermal reservoir. But, in contrast to the state probabilities in a closed system, each state does *not* have the same probability. That is, the system does not spend the same fraction of time in each state. The key to the canonical formalism is the determination of the probability distribution of the system among its microstates. And this problem is solved by the realization that the system plus the reservoir constitute a *closed* system, to which the principle of equal probability of microstates again applies.

A simple analogy is instructive. Consider a set of three dice, one of which is red (the remaining two being white). The three dice have been "thrown" many thousands of times. Whenever the sum of the numbers on the three dice has been 12 (and only then), the number on the red die has been recorded. In what fraction of these recorded throws has the red die shown a one, a two, . . . , a six?

The result, left to the reader, is that the red die has shown a one in $\frac{2}{25}$ of the throws, a two in $\frac{3}{25}, \ldots$, a five in $\frac{6}{25}$, and a six in $\frac{5}{25}$ of the recorded throws. The probability of a (red) six, in this restricted set of throws, is $\frac{1}{5}$.

The red die is the analogue of our system of interest, the white dice correspond to the reservoir, the numbers shown correspond to the energies of the respective systems, and the restriction to throws in which the sum is 12 corresponds to the constancy of the total energy (of system plus reservoir).

The probability f_j of the subsystem being in state j is equal to the fraction of the total number of states (of system-plus-reservoir) in which the subsystem is in the state j (with energy E_j):

$$f_j = \frac{\Omega_{\text{res}}(E_{\text{tot}} - E_j)}{\Omega_{\text{tot}}(E_{\text{tot}})} \tag{16.1}$$

Here E_{tot} is the total energy of the system-plus-reservoir, and Ω_{tot} is the total number of states of the system-plus-reservoir. The quantity in the numerator, $\Omega_{\text{res}}(E_{\text{tot}} - E_j)$ is the number of states available to the reservoir when the subsystem is in the state j (leaving energy $E_{\text{tot}} - E_j$ in the reservoir).

This is the seminal relation in the canonical formalism, but it can be re-expressed in a far more convenient form. The denominator is related to the entropy of the composite system by equation 15.1. The numerator is similarly related to the entropy of the reservoir, so that

$$f_j = \frac{\exp\{k_B^{-1} S_{\text{res}}(E_{\text{tot}} - E_j)\}}{\exp\{k_B^{-1} S_{\text{tot}}(E_{\text{tot}})\}} \tag{16.2}$$

If U is the average value of the energy of the subsystem, then the additivity of the entropy implies

$$S_{\text{tot}}(E_{\text{tot}}) = S(U) + S_{\text{res}}(E_{\text{tot}} - U) \tag{16.3}$$

Furthermore, expanding $S_{\text{res}}(E_{\text{tot}} - E_j)$ around the equilibrium point $E_{\text{tot}} - U$;

$$\begin{aligned} S_{\text{res}}(E_{\text{tot}} - E_j) &= S_{\text{res}}(E_{\text{tot}} - U + U - E_j) \\ &= S_{\text{res}}(E_{\text{tot}} - U) + (U - E_j)/T \end{aligned} \tag{16.4}$$

No additional terms in the expansion appear (this being the very definition of a reservoir). Inserting these latter two equations in the expression for f_j,

$$f_j = e^{(1/k_B T)\{U - TS(U)\}} e^{-(1/k_B T) E_j} \tag{16.5}$$

The quantity $1/k_BT$ appears so pervasively throughout the theory that it is standard practice to adopt the notation

$$\beta \equiv 1/(k_B T) \tag{16.6}$$

Furthermore $U - TS(U)$ is the Helmholtz potential of the system, so that we finally achieve the fundamental result for the probability f_j of the subsystem being in the state j

$$f_j = e^{\beta F} e^{-\beta E_j} \tag{16.7}$$

Of course the Helmholtz potential is not known; it is in fact our task to compute it. The key to its evaluation is the observation that $e^{\beta F}$ plays the role of a state-independent normalization factor in equation 16.7.

$$\sum_j f_j = e^{\beta F} \sum_j e^{-\beta E_j} = 1 \tag{16.8}$$

or

$$e^{-\beta F} = Z \tag{16.9}$$

where Z, the "canonical partition sum," is defined by

$$Z \equiv \sum_j e^{-\beta E_j} \tag{16.10}$$

We have now formulated a complete algorithm for the calculation of a fundamental relation in the canonical formalism. Given a list of all states j of the system, and their energies E_j, we calculate the partition sum (16.10). *The partition sum is thus obtained as a function of temperature (or β) and of the parameters ($V, N_1, N_2, \ldots$) that determine the energy levels. Equation* 16.9 *in turn determines the Helmholtz potential as a function also of T, V, N_1, N_r. This is the sought for fundamental relation.*

The entire algorithm is summarized in the relation

$$-\beta F = \ln \sum_j e^{-\beta E_j} \equiv \ln Z$$

which should be committed to memory.

A corroboration of the consistency of the formalism follows from recalling that f_j is the probability of occupation of the jth state, which (from equations 16.7, 16.9 and 16.10) can be written in the very useful form

$$f_j = e^{-\beta E_j} \Big/ \sum_i e^{-\beta E_i} \tag{16.11}$$

The average energy is then expected to be

$$U = \sum_j E_j f_j = \sum_j E_j e^{-\beta E_j} \Big/ \sum_i e^{-\beta E_i} \tag{16.12}$$

or

$$U = -(d/d\beta) \ln Z \tag{16.13}$$

Insertion of equation 16.9, expressing Z in terms of F, and recalling that $\beta = 1/k_B T$ reduces this equation to the familiar thermodynamic relation $U = F + TS = F - T(\partial F/\partial T)$ and thereby confirms its validity. Equations 16.12 and 16.13 are very useful in statistical mechanics, but it must be stressed that these equations do not constitute a fundamental relation. The fundamental relation is given by equations 16.9 and 16.10, giving F (rather than U) as a function of β, V, N.

A final observation on units and on formal structure is revealing. The quantity β is, of course, merely the reciprocal temperature in "natural units." The canonical formalism then gives the quantity βF in terms of β, V, and N. That is, F/T is given as a function of $1/T$, V, and N. *This is a fundamental equation in the representation* $S[1/T]$ (recall Section 5.4). Just as the microcanonical formalism is naturally expressed in entropy representation, the canonical formalism is naturally expressed in $S[\beta]$ representation. The generalized canonical representations to be discussed in Chapter 17 will similarly all be expressed most naturally in terms of Massieu functions. Nevertheless we shall conform to universal usage and refer to the canonical formalism as being based on the Helmholtz potential. No formal difficulties arise from this slight "misrepresentation."

PROBLEMS

16.1-1. Show that equation 16.13 is equivalent to $U = F + TS$.

16.1-2. From the canonical algorithm expressed by equations 16.9 and 16.10, express the pressure in terms of a derivative of the partition sum. Further, express the pressure in terms of the derivatives $\partial E_j/\partial V$ (and of T and the E_j). Can you give a heuristic interpretation of this equation?

16.1-3. Show that $S/k_B = \beta^2 \partial F/\partial\beta$ and thereby express S in terms of Z and its derivatives (with respect to β).

16.1-4. Show that $c_v = -\beta(\partial s/\partial\beta)_v$ and thereby express c_v in terms of the partition sum and its derivatives (with respect to β).

Answer:

$$c_v = N^{-1} k_B \beta^2 \frac{\partial^2 \ln Z}{\partial \beta^2}$$

16-2 ADDITIVE ENERGIES AND FACTORIZABILITY OF THE PARTITION SUM

To illustrate the remarkable simplicity of the canonical formalism we recall the two-state system of Section 15.3. In that model $\tilde{N}$ distinguishable "atoms" each were presumed to have two permissible states, of energies 0 and ε. Had we attributed even only three states to each atom the problem would have become so difficult as to be insoluble by the microcanonical formalism, at least for general values of the excitation energies. By the canonical formalism it is simple indeed!

We consider a system composed of $\tilde{N}$ distinguishable "elements," an element being an independent (noninteracting) excitation mode of the system. If the system is composed of noninteracting material constituents, such as the molecules of an ideal gas, the "elements" refer to the excitations of the individual molecules. In strongly interacting systems the elements may be wavelike collective excitations such as vibrational modes or electromagnetic modes. *The identifying characteristic of an "element" is that the energy of the system is a sum over the energies of the elements, which are independent and noninteracting.*

Each element can exist in a set of *orbital states* (we henceforth use the term *orbital state* to distinguish the states of an element from the states of the collective system). The energy of the ith element in its jth orbital state is ε_{ij}. Each of the elements need not be the same, either in the energies or the number of its possible orbital states. *The total energy of the system is the sum of the single-element energies, and each element is permitted to occupy any one of its orbital states independently of the orbital states of the other elements*. Then the partition sum is

$$Z = \sum_{j, j', j'', \dots} e^{-\beta(\varepsilon_{1j} + \varepsilon_{2j'} + \varepsilon_{3j''} + \cdots)} \tag{16.14}$$

$$= \sum_{j, j', j'', \dots} e^{-\beta\varepsilon_{1j}} e^{-\beta\varepsilon_{2j'}} e^{-\beta\varepsilon_{3j''}} \cdots \tag{16.15}$$

$$= \sum_{j} e^{-\beta\varepsilon_{1j}} \sum_{j'} e^{-\beta\varepsilon_{2j'}} \sum_{j''} e^{-\beta\varepsilon_{3j''}} \cdots \tag{16.16}$$

$$= z_1 z_2 z_3 \cdots \tag{16.17}$$

where z_i, the "partition sum of the ith element," is

$$z_i = \sum_{j} e^{-\beta\varepsilon_{ij}} \tag{16.18}$$

The partition sum factors. Furthermore the Helmholtz potential is additive *over elements*

$$-\beta F = \ln Z = \ln z_1 + \ln z_2 + \cdots \tag{16.19}$$

This result is so remarkably simple, powerful, and useful that we emphasize again that it applies to any system in which (a) the energy is additive over elements and (b) each element is permitted to occupy any of its orbital states independently of the orbital state of any other element.

The "two-state model" of Section 15.3 satisfies the above criteria, whence

$$Z = z^{\tilde{N}} = (1 + e^{-\beta\varepsilon})^{\tilde{N}} \tag{16.20}$$

and

$$F = -\tilde{N} k_B T \ln(1 + e^{-\beta\varepsilon}) \tag{16.21}$$

It is left to the reader to demonstrate that this solution is equivalent to that found in Section 15.3. If the number of orbitals had been three rather than two, the partition sum per particle z would merely have contained three terms and the Helmholtz potential would have contained an additional term in the argument of the logarithm.

The Einstein model of a crystal (Section 15.2) similarly yields to the simplicity of the canonical formalism. Here the "elements" are the vibrational modes, and the partition sum per mode is

$$z = 1 + e^{-\beta\hbar\omega_0} + e^{-2\beta\hbar\omega_0} + \cdots = \sum_{n=0}^{\infty} e^{-n\beta\hbar\omega_0} \tag{16.22}$$

This "geometric series" sums directly to

$$z = \frac{1}{1 - e^{-\beta\hbar\omega_0}} \tag{16.23}$$

There are $3\tilde{N}$ vibrational modes so that the fundamental equation of the Einstein model, in the canonical formalism, is

$$F = -\beta^{-1} \ln z^{3\tilde{N}} = 3\tilde{N} k_B T \ln(1 - e^{-\beta\hbar\omega_0}) \tag{16.24}$$

Clearly Einstein's drastic assumption that all modes of vibration of the crystal have the same frequency is no longer necessary in this formalism. A more physically reasonable approximation, due to P. Debye, will be discussed in Section 16.7.

PROBLEMS

16.2-1. Consider a system of three particles, each different. The first particle has two orbital states, of energies ε_{11} and ε_{12}. The second particle has permissible energies ε_{21} and ε_{22}, and the third particle has permissible energies ε_{31} and ε_{32}. Write the partition sum explicitly in the form of equation 16.14, and by explicit algebra, factor it in the form of equation 16.17.

16.2-2. Show that for the two-level system the Helmholtz potential calculated in equation 16.21 is equivalent to the fundamental equation found in Section 15.3.

16.2-3. Is the energy additive over the particles of a gas if the particles are uncharged mass points (with negligible gravitational interaction)? Is the partition sum factorizable if half the particles carry a positive electric charge and half carry a negative electric charge? Is the partition sum factorizable if the particles are "fermions" obeying the Pauli exclusion principle (such as neutrinos)?

16.2-4. Calculate the heat capacity per mode from the fundamental equation 16.24.

16.2-5. Calculate the energy per mode from equation 16.24. What is the leading term in $U(T)$ in the regions of $T \simeq 0$ and of T large?

16.2-6. A binary alloy is composed of $\tilde{N}_A$ atoms of type A and of $\tilde{N}_B$ atoms of type B. Each A-type atom can exist in its ground state or in an excited state of energy ε (all other states are of such high energy that they can be neglected at the temperatures of interest). Each B-type atom similarly can exist in its ground state of energy zero or in an excited state of energy 2ε. The system is in equilibrium at temperature T.

a) Calculate the Helmholtz potential of the system.

b) Calculate the heat capacity of the system.

16.2-7. A paramagnetic salt is composed of 1 mole of noninteracting ions, each with a magnetic moment of one Bohr magneton ($\mu_B = 9.274 \times 10^{-24}$ joules/tesla). A magnetic field B_e is applied along a particular direction; the permissible states of the ionic moments are either parallel or antiparallel to this direction.

a) Assuming the system is maintained at a temperature $T = 4$ K and B_e is increased from 1 Tesla to 10 Tesla, what is the magnitude of the heat transfer from the thermal reservoir?

b) If the system is now thermally isolated and the applied magnetic field B_e is decreased from 10 Tesla to 1 Tesla, what is the final temperature of the system? (This process is referred to as cooling by adiabatic demagnetization.)

16-3 INTERNAL MODES IN A GAS

The excitations of the molecules of a gas include the three translational modes of the molecules as a whole, vibrational modes, rotational modes, electronic modes, and modes of excitation of the nucleus. For simplicity we initially assume that each of these modes is independent, later returning to reexamine this assumption. Then the partition sum factors with respect to the various modes

$$Z = Z_{\text{trans}} Z_{\text{vib}} Z_{\text{rot}} Z_{\text{elect}} Z_{\text{nuc}} \tag{16.25}$$

and, further, with respect to the molecules

$$Z_{\text{vib}} = z_{\text{vib}}^{\tilde{N}}, \qquad Z_{\text{rot}} = z_{\text{rot}}^{\tilde{N}} \tag{16.26}$$

and similarly for Z_{elect} and Z_{nuc}.

The "ideality" or "nonideality" of the gas is a property primarily of the translational partition sum. The translational modes in any case warrant a separate and careful treatment, which we postpone to Section 16.10. We now simply assume that any intermolecular collisions do not couple to the internal modes (rotation, vibration, etc.).

The $\tilde{N}$ identical vibrational modes of a given type (one centered on each molecule) are formally identical to the vibrational modes of the Einstein model of a crystal; that is, they are just simple harmonic oscillators. For a mode of frequency ω_0

$$Z_{\text{vib}} = z_{\text{vib}}^{\tilde{N}} = \left(1 - e^{-\beta\hbar\omega_0}\right)^{-\tilde{N}} \tag{16.27}$$

and the contribution of this vibrational mode to the Helmholtz potential is as given in equation 16.24 (with $3\tilde{N}$ replaced by $\tilde{N}$). The contribution of a vibrational mode to the heat capacity of the gas is then as shown in Fig. 15.2 (the ordinate being c/R rather than $c/3R$). As described in Section 13.1, the heat capacity "rises in a roughly steplike fashion" in the vicinity of $k_B T \simeq \hbar\omega_0$, and it asymptotes to $c = R$. Figure 13.1 was plotted as the sum of contributions from two vibrational modes, with $\omega_2 = 15\omega_1$.

The characteristic vibrational temperature $\hbar\omega_0/k_B$ ranges from several thousand kelvin for molecules containing very light elements ($\simeq 6300$ K for H_2) to several hundred kelvin for molecules containing heavier elements ($\simeq 309$ K for Br_2).

To consider the rotational modes of a gas we focus particularly on heteronuclear diatomic molecules (such as HCl), which require two angular coordinates to specify their orientation. The rotational energy of such heteronuclear diatomic molecules is quantized, with energy eigenvalues given by

$$\varepsilon_{\ell} = \ell(\ell + 1)\varepsilon \qquad \ell = 0, 1, 2, \ldots \tag{16.28}$$

Each energy level is $(2\ell + 1)$-fold degenerate. The energy unit ε is equal to $\frac{1}{2}\hbar^2/(\text{moment of inertia})$, or approximately 2×10^{-21} J for the HCl molecule. The characteristic separation between levels is of the order of ε, which corresponds to a temperature $\varepsilon/k_B \simeq 15$ K for HCl—larger for lighter molecules and smaller for heavier molecules.

The rotational partition sum per molecule is

$$z_{\text{rot}} = \sum_{\ell=0}^{\infty} (2\ell + 1) e^{-\beta\ell(\ell+1)\varepsilon} \tag{16.29}$$

If $k_B T \gg \varepsilon$ the sum can be replaced by an integral. Then, noting that $2\ell + 1$ is the derivative of $\ell(\ell + 1)$, and writing x for the quantity $\ell(\ell+1)$,

$$z_{\text{rot}} \simeq \int_0^{\infty} e^{-\beta\varepsilon x}\, dx = \frac{1}{\beta\varepsilon} = \frac{k_B T}{\varepsilon} \tag{16.30}$$

If $k_B T$ is less than or of the order of ε it may be practical to calculate several terms of the series explicitly, to some ℓ' such that $\ell'(\ell' + 1) \gg k_B T$, and to integrate over the remaining range (from ℓ' to infinity); see Problem 16.3-2.

It is left to the reader to show that for $k_B T \gg \varepsilon$ the average energy is $k_B T$.

The case of homonuclear diatomic molecules, such as O_2 or H_2, is subject to quantum mechanical symmetry conditions into which we shall not enter. Only the even terms in the partition sum, or only the odd terms, are permitted (depending upon detailed characteristics of the atoms). At high temperatures this restriction merely halves the rotational partition sum per molecule.

The nuclear and electronic contributions can be computed in similar fashion, but generally only the lowest energy levels of each contribute. Then z_{nuc} is simply the "degeneracy" (multiplicity) of the lowest energy configuration. Each of these factors simply contributes $\tilde{N} k_B T \ln$ (multiplicity) to the Helmholtz potential.

It is of interest to return to the assumption that the various modes are independent. This assumption is generally a good (but *not* a rigorous) approximation. Thus the vibrations of a diatomic molecule change the instantaneous interatomic distance and thereby change the instantaneous moment of inertia of rotation. It is only because the vibrations generally are very fast relative to the rotations that the rotations sense only the *average* interatomic distance, and thereby become effectively independent of the vibrations.