Chapter 4
Various Statistical Ensembles

A canonical ensemble describes a system held at a constant temperature. From the canonical partition function follows the Helmholtz free energy. As we saw in thermodynamics, however, it is highly desirable to be able to describe a system held under a different set of constraints, such as constant temperature and pressure or constant temperature and chemical potentials of some species. These situations call for different free energies in thermodynamics, to which correspond different statistical ensembles in statistical mechanics. In this chapter, we construct such ensembles and illustrate their applications with several simple examples. Our first task is to establish the notion of microcanonical ensemble, which is suitable for describing an isolated system. The other ensembles of more practical importance, including the canonical ensemble, can be derived straightforwardly from the microcanonical ensemble.

4.1 Fluctuations in a Canonical Ensemble

A system held at a constant temperature is constantly exchanging energy with the surroundings, which serves as a heat bath. As a result, the energy of the system will fluctuate with time. Let us first figure out the magnitude of this fluctuation. A good measure is given by

$$\mathit{\Delta}_{\mathrm{rms}} H := \left\langle (H - \langle H \rangle)^2 \right\rangle^{1/2} . \tag{4.1}$$

In fact, $\mathit{\Delta}_{\mathrm{rms}} H$ is zero if and only if H is identically equal to $\langle H \rangle$, and increases as the deviation of H from the average becomes more probable. Because of the square root, $\mathit{\Delta}_{\mathrm{rms}} H$ also has the dimension of energy.

Recall from Exercise 3.2 that

$$\langle H \rangle = -\frac{\partial \ln Z}{\partial \beta} \quad \text{and} \quad \left\langle (H - \langle H \rangle)^2 \right\rangle = \frac{\partial^2 \ln Z}{\partial \beta^2} . \tag{4.2}$$

I. Kusaka, *Statistical Mechanics for Engineers*,
DOI 10.1007/978-3-319-13809-1_4

For an ideal gas of N noninteracting identical particles, we have

$$Z = \frac{V^N}{N!\Lambda^{3N}} , \tag{4.3}$$

where $\Lambda \sim \sqrt{\beta}$. Thus,

$$\langle H \rangle = \frac{3N}{2\beta}, \quad \left\langle (H - \langle H \rangle)^2 \right\rangle = \frac{3N}{2\beta^2} , \tag{4.4}$$

and hence

$$\frac{\Delta_{\text{rms}} H}{\langle H \rangle} \sim \sqrt{\frac{1}{N}} . \tag{4.5}$$

For a macroscopic system of $N \sim 10^{24}$, this quantity is 10^{-12}. If we let $\langle H \rangle = 1$ J for the sake of illustration, then the magnitude of fluctuation $\Delta_{\text{rms}} H$ is of the order of 10^{-12} J. Such an extremely small quantity cannot be measured by an instrument whose intended range is of the order of 1 J.

Equations similar to (4.5) hold quite generally for any extensive quantities of a homogeneous macroscopic body. To see this, let G be such an extensive quantity and denote by N the total number of particles in the body. We divide the body into M equal parts and denote the value of the extensive quantity pertaining to the ith part by g_i. Since G is an extensive quantity, we may write

$$G = \sum_{i=1}^{M} g_i . \tag{4.6}$$

In writing this equation, we are actually assuming that each part is still sufficiently large that the contribution to G from the interaction among adjacent parts can be ignored.[19] The assumption is justified since we are interested only in macroscopic systems here and N can be taken as large as we desire. Taking the average of (4.6), we find

$$\langle G \rangle = \sum_{i=1}^{M} \langle g_i \rangle = \sum_{i=1}^{M} g = Mg , \tag{4.7}$$

where we used the fact that the original system is homogeneous and hence $g := \langle g_i \rangle$ is common for all i. Let us next figure out the magnitude of fluctuation of G. First, we calculate

$$(G - \langle G \rangle)^2 = \left(\sum_{i=1}^{M} g_i - Mg \right)^2 = \left[\sum_{i=1}^{M} (g_i - g) \right]^2 , \tag{4.8}$$

where we used (4.7). Now,

$$\left[\sum_{i=1}^{M} (g_i - g) \right]^2 = \sum_{i=1}^{M} (g_i - g) \sum_{j=1}^{M} (g_j - g) = \sum_{i=1}^{M} \sum_{j=1}^{M} (g_i - g)(g_j - g) . \tag{4.9}$$

Taking the average, we find

$$(\Delta_{\mathrm{rms}}G)^2 = \sum_{i=1}^{M}\sum_{j=1}^{M}\langle (g_i - g)(g_j - g)\rangle \,. \tag{4.10}$$

We have already assumed that the interaction among the adjacent parts is sufficiently weak, implying that we may regard each part as statistically independent of the adjacent ones. Thus, if $i \neq j$,

$$\langle (g_i - g)(g_j - g)\rangle = \langle g_i - g\rangle\langle g_j - g\rangle = 0 \,. \tag{4.11}$$

However, if $i = j$,

$$\langle (g_i - g)(g_j - g)\rangle = \langle (g_i - g)^2\rangle =: (\Delta_{\mathrm{rms}}\, g)^2 \tag{4.12}$$

and (4.10) becomes

$$(\Delta_{\mathrm{rms}}G)^2 = M(\Delta_{\mathrm{rms}}\, g)^2 \,. \tag{4.13}$$

Combining (4.7) and (4.13), we arrive at

$$\frac{\Delta_{\mathrm{rms}}G}{\langle G\rangle} \sim \sqrt{\frac{1}{M}} \,. \tag{4.14}$$

If we fix the number of particles included in each of the parts, into which we divided our original system, then $M \sim N$, and hence

$$\frac{\Delta_{\mathrm{rms}}G}{\langle G\rangle} \sim \sqrt{\frac{1}{N}} \tag{4.15}$$

as advertised.

4.2 Microcanonical Ensemble

Equation (4.5) indicates that the distribution of the system energy is sharply peaked around $\langle H\rangle$ for macroscopic systems with $N \sim 10^{24}$. This being the case, the canonical ensemble can be approximated by a new ensemble, called **microcanonical ensemble**, in which all the copies have practically the same energy.

4.2.1 Expression for ρ

To define a statistical ensemble precisely, we must give an explicit expression for ρ. For this purpose, let us start from a canonical ensemble and then imagine selecting from the ensemble only those copies whose energy lies within an extremely narrow semi-open interval $(E-\Delta E,E]$. (That is, H of a copy must satisfy $E-\Delta E<H\le E$.) Those copies so selected form the microcanonical ensemble. Now, we recall that copies in the original canonical ensemble are distributed in the phase space according to the density ρ which depends only on the energy of the system. Since $\Delta E/E$ is extremely small, we can regard ρ within the interval $(E-\Delta E,E]$ as constant without any noticeable loss of accuracy. The microcanonical ensemble constructed from the canonical ensemble is therefore characterized by

$$\rho(q^f,p^f)\mathrm{d}q^f\mathrm{d}p^f=\begin{cases}\frac{1}{C_M}\frac{\mathrm{d}q^f\mathrm{d}p^f}{h^f\mathscr{P}} & \text{if } E-\Delta E<H(q^f,p^f)\le E\\ 0 & \text{otherwise.}\end{cases} \tag{4.16}$$

Aside from the normalization factor, this expression for ρ when $E-\Delta E<H\le E$ follows from (3.165) by replacing the Boltzmann factor $e^{-\beta H}$ by a constant.

There is nothing special about our choice of the interval for H. We could have very well chosen $[E-\Delta E,E)$ instead. The particular choice we made here, however, will affect the definition of $\mathscr{W}(E)$ and our choice for the step function to be introduced in what follows. The reason for using a half-open interval becomes clear in (4.23).

If we take a phase point A, the number of copies in the ensemble whose representative phase points fall inside the volume element $\mathrm{d}q^f\mathrm{d}p^f$ taken around A is given by $\mathscr{N}\rho(q^f,p^f)\mathrm{d}q^f\mathrm{d}p^f$, where $\mathscr{N}$ is the total number of copies in the ensemble. According to (4.16), if the $\mathrm{d}q^f\mathrm{d}p^f$ lies entirely inside the region $E-\Delta E<H(q^f,p^f)\le E$, this number is $\mathscr{N}\mathrm{d}q^f\mathrm{d}p^f/h^f\mathscr{P}C_M$ regardless of exactly where A is located. On the other hand, if $\mathrm{d}q^f\mathrm{d}p^f$ is taken entirely outside the indicated region, the number is strictly zero. Thus, the copies in the ensemble are distributed with uniform density in the region of the phase space corresponding to $(E-\Delta E,E]$. In other words, all states satisfying $E-\Delta E<H(q^f,p^f)\le E$ are equally probable. This result is known as the **principle of equal weight**.

We arrived at this principle by applying a canonical ensemble to a macroscopic system. In a more common approach to statistical mechanics, one starts with a microcanonical ensemble. In this case, the principle plays a more central role of being the logical foundation of statistical mechanics.

We recall that ρ must be normalized:

$$\int\rho(q^f,p^f)\mathrm{d}q^f\mathrm{d}p^f=1\,. \tag{4.17}$$

Substituting (4.16), we find

$$C_M=\frac{1}{h^f\mathscr{P}}\int_{E-\Delta E<H\le E}\mathrm{d}q^f\mathrm{d}p^f\,. \tag{4.18}$$

We refer to C_M as the **microcanonical partition function**. Note that the integral in (4.18) is the volume of the region in the phase space (or the phase volume for short) that is compatible with $E-\Delta E<H\le E$, while $h^f\mathscr{P}$ is the phase volume occupied by a single (quantum mechanically distinct) state. Thus, C_M is the number of states consistent with the condition that $E-\Delta E<H\le E$.

We introduced a microcanonical ensemble as an approximation to a canonical ensemble. Thus, it is sensible to expect that Gibbs's entropy formula continues to apply to the microcanonical ensemble. Using (4.16) in (3.168) and noting that $x\ln x \to 0$ as $x \to 0$,[20] we have

$$-S/k_B = \int \rho \ln(h^f \mathscr{P} \rho) \mathrm{d}q^f \mathrm{d}p^f = \int \rho \ln \frac{1}{C_M} \mathrm{d}q^f \mathrm{d}p^f = \ln \frac{1}{C_M} \ . \tag{4.19}$$

The last step follows from (4.17) and the fact that C_M is a constant. This is the famous **Boltzmann's entropy formula**:

$$S = k_B \ln C_M \ . \tag{4.20}$$

4.2.2 *Choice of ΔE*

It is clear from (4.18) that, once the functional form of H is specified and the integration is carried out, the resulting C_M and hence S depend on f, E, ΔE, and the limits of integrations. Because the integration in (4.18) includes only those microstates compatible with $E - \Delta E < H \leq E$ and $\Delta E/E$ is extremely small, the internal energy may be identified with E:

$$U := \langle H \rangle \approx E \ . \tag{4.21}$$

For a system of N particles contained in a three-dimensional box, $f = 3N$ and the limits of integrations for their coordinates usually show up only in the form of the system volume V in the final expression for C_M. This being the case,

$$S = S(U, V, N, \Delta E) \ . \tag{4.22}$$

Of all these quantities, upon which S can depend, ΔE alone does not have any corresponding quantity in thermodynamics. We now show that the value of ΔE can be taken quite arbitrarily within a reasonable bound with no quantitative consequence to the value of S. Thus, S is practically a function only of U, V, and N.

To see that S is in fact insensitive to ΔE, let $\mathscr{W}(E, V, N)$ denote the number of states satisfying $H \leq E$ for given values of V and N. Clearly,

$$C_M(E, V, N, \Delta E) = \mathscr{W}(E, V, N) - \mathscr{W}(E - \Delta E, V, N) \ . \tag{4.23}$$

Because ΔE is extremely small compared to E, we expand the second $\mathscr{W}$ into a Taylor series to obtain

$$\begin{aligned} C_M(E, V, N, \Delta E) &\approx \mathscr{W}(E, V, N) - \left[\mathscr{W}(E, V, N) - \frac{\partial \mathscr{W}}{\partial E} \Delta E \right] \\ &= \frac{\partial \mathscr{W}}{\partial E} \Delta E =: \overline{\Omega}(E, V, N) \Delta E \ , \end{aligned} \tag{4.24}$$

where

$$\overline{\Omega}(E,V,N) := \frac{\partial \mathscr{W}}{\partial E} \tag{4.25}$$

is the **density of states**. This terminology is quite appropriate since an integration of $\overline{\Omega}$ over a certain interval of energy yields the total number of states whose energy falls within that interval.

Using (4.20), we obtain

$$S = k_B \ln \overline{\Omega}(E,V,N)\Delta E \ , \tag{4.26}$$

which is still a function of E, V, N, and ΔE.

Based on the consideration that led us to microcanonical ensemble, it seems quite reasonable to set $\Delta E = E/\sqrt{N}$, for which the entropy is evaluated as

$$S_1 = k_B \ln \overline{\Omega}(E,V,N)\frac{E}{\sqrt{N}} \ . \tag{4.27}$$

An extreme choice for ΔE would be E itself. For this choice, the Taylor series expansion used in arriving at (4.24) cannot be justified. Nevertheless, if we proceed blindly, we find

$$S_2 = k_B \ln \overline{\Omega}(E,V,N)E \ . \tag{4.28}$$

Thus,

$$\frac{S_2 - S_1}{k_B N} = \frac{1}{2N}\ln N \ . \tag{4.29}$$

For a macroscopic body with $N \sim 10^{24}$, this quantity is approximately 3×10^{-23} even though the two choices for ΔE differ by a factor of $\sqrt{N} \sim 10^{12}$. For macroscopic systems, therefore, the precise value of ΔE is unimportant. What happens if N is not large enough? In this case, the use of microcanonical ensemble is of little interest. The ensemble was introduced only as an approximation applicable for a large N.

4.2.3 Isolated System

In a microcanonical ensemble, all copies of the ensemble have nearly the same amount of energy. If we recall how ρ was constructed from the long-time behavior of a single system, this implies that its energy fluctuates very little. In fact, the width of this fluctuation relative to the energy itself is of the order of $1/\sqrt{N}$ and E of the system is a constant within the accuracy of any practical means of measuring E. We recall that the energy of an isolated system is a constant of motion. Thus, a microcanonical ensemble is a natural choice for describing a system that can be regarded as isolated within the accuracy of experiments.

This correspondence between a microcanonical ensemble and an isolated system is quite satisfactory from the point of view of thermodynamics. We recall that entropy played a prominent role in the condition of equilibrium of an isolated system. Moreover, S expressed in terms of (U,V,N) is a fundamental equation of thermodynamics. Because S is practically independent of ΔE, (4.22) indicates that S is obtained directly as a function of just these variables.

One might argue that ΔE should be identically zero for a truly isolated system. While this might be so, the system we intend to study must be subjected to experimental measurements by necessity. Because ΔE is extremely small compared to E itself, the very attempt to measure something about the system will generally affect the system energy by an amount far exceeding ΔE. At the same time, we have no interest in predicting the behavior of a system if no measurement will ever be performed on it. In this sense, a truly isolated system never really arises as a subject of our study. Instead, ΔE is nonzero for isolated systems of our interest.

4.3 Phase Integral in Microcanonical Ensemble

The phase integral for C_M extends only over a portion of the phase space consistent with the condition $E - \Delta E < H \leq E$ as we see from (4.18). It is very awkward to work with a phase integral whose limits are given in terms of the system energy. For the subsequent development, it is desirable, and indeed possible, to rewrite the phase integral so that the integral is over all phase space.

We first note that the phase integral

$$\int_{H \leq E} \mathrm{d}q^f \mathrm{d}p^f \tag{4.30}$$

is the phase volume of the region satisfying $H(q^f, p^f) \leq E$. If we divide this quantity by $h^f \mathscr{P}$, which represents the phase volume occupied by a single microstate, we obtain the number of states with $H(q^f, p^f) \leq E$. That is,

$$\mathscr{W}(E,V,N) = \frac{1}{h^f \mathscr{P}} \int_{H \leq E} \mathrm{d}q^f \mathrm{d}p^f = \frac{1}{h^f \mathscr{P}} \int \theta \left(E - H(q^f, p^f)\right) \mathrm{d}q^f \mathrm{d}p^f \,, \tag{4.31}$$

where θ is the step function defined by (D.2). See Fig. 4.1 to justify its use here.

Following the definition of $\overline{\Omega}(E)$ given by (4.24), we now have to differentiate $\mathscr{W}$ with respect to E. Because the integral is taken over all phase space, the only E dependence of $\mathscr{W}$ is in the step function θ, giving rise to the Dirac δ-function as shown in Appendix D:

$$\begin{aligned} \overline{\Omega}(E,V,N) = \frac{\partial \mathscr{W}}{\partial E} &= \frac{1}{h^f \mathscr{P}} \int \frac{\partial}{\partial E} \theta \left(E - H(q^f, p^f)\right) \mathrm{d}q^f \mathrm{d}p^f \\ &= \frac{1}{h^f \mathscr{P}} \int \delta \left(E - H(q^f, p^f)\right) \mathrm{d}q^f \mathrm{d}p^f \,. \end{aligned} \tag{4.32}$$

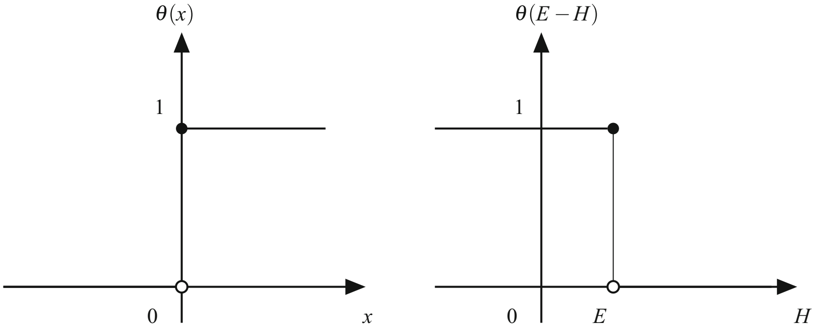


Fig. 4.1 Step function. If $H > E$, $E - H < 0$ and hence $\theta(E-H) = 0$. If $H \le E$, $E - H \ge 0$ and $\theta(E-H) = 1$.

Substituting this expression into (4.24), we find that

$$C_M = \frac{1}{h^f \mathscr{P}} \int \delta\left(E - H(q^f, p^f)\right) \Delta E \mathrm{d}q^f \mathrm{d}p^f \ . \tag{4.33}$$

In essence, we replaced the limits of the phase integral in (4.18) by a δ-function in the integrand.

Example 4.1. Irreversible expansion of an ideal gas: Consider an isolated system consisting of two compartments of equal volume V that are separated by rigid and impermeable partition held fixed in place. Suppose that, at the initial state, one of the compartments is filled with N noninteracting identical particles while the other compartment is empty. If the partition is removed, the gas will expand and fill the entire volume $2V$. Insofar as this process is irreversible, we expect that the entropy of the system at the final state is larger than that at the initial state. Let us compute the change in entropy by means of (4.33).

At the initial state, the Hamiltonian of the system is given by

$$H(\boldsymbol{r}^N, \boldsymbol{p}^N) = \begin{cases} \sum_{i=1}^N |\boldsymbol{p}_i|^2 / 2m & \text{if } \boldsymbol{r}_i \in V \text{ for all } i \\ \infty & \text{otherwise.} \end{cases} \tag{4.34}$$

Setting $\mathscr{P} = N!$ and $f = 3N$ we write (4.33) as

$$C_M^i = \frac{1}{N! h^{3N}} \int \delta\left(E - H(\boldsymbol{r}^N, \boldsymbol{p}^N)\right) \Delta E \mathrm{d}\boldsymbol{r}^N \mathrm{d}\boldsymbol{p}^N \ . \tag{4.35}$$

Because we are concerned only with a finite value of E, the integrand is zero if H is infinity, that is, if any one of the particle goes outside the volume V. It follows that the integration over coordinates should be confined to the volume

V and H is then a function of momenta only. Thus,

$$C_M^i = \frac{1}{N!h^{3N}} \int_{\boldsymbol{r}^N \in V} \left[\int \delta\left(E - H(\boldsymbol{p}^N)\right) \Delta E \mathrm{d}\boldsymbol{p}^N \right] \mathrm{d}\boldsymbol{r}^N , \tag{4.36}$$

where $\boldsymbol{r}^N \in V$ denotes the condition that all the particles are inside V. We have also chosen to perform first the integration with respect to momenta. But, the result of this integration is independent of coordinates. So,

$$C_M^i = \frac{V^N}{N!h^{3N}} \int \delta\left(E - H(\boldsymbol{p}^N)\right) \Delta E \mathrm{d}\boldsymbol{p}^N . \tag{4.37}$$

In this problem, there is no need to compute this integral as we will see shortly. Nevertheless, such a simple system as an ideal gas should surely be "easy" to handle and you are invited to carry out the actual computation in Exercise 4.3.

At the final state, particles can be anywhere inside the volume $2V$. Repeating the above computation, but with V replaced by $2V$, we find

$$C_M^f = \frac{(2V)^N}{N!h^{3N}} \int \delta\left(E - H(\boldsymbol{p}^N)\right) \Delta E \mathrm{d}\boldsymbol{p}^N . \tag{4.38}$$

The change in entropy upon removal of the partition, therefore, is given by

$$S^f - S^i = k_B \ln C_M^f - k_B \ln C_M^i = k_B \ln \frac{C_M^f}{C_M^i} = Nk_B \ln 2 , \tag{4.39}$$

which is positive as expected.

Note that C_M^f is the number of distinct microstates consistent with the given values of E, N, and the volume $2V$. We also recall that all these C_M^f states are equally probable. Among these states, however, some have all the particles confined to the original volume V. The number of such states is just C_M^i.

(To see this more clearly, suppose that the system is a one-dimensional box of length $2L$ and let $N = 2$. The integration over coordinates is then given by

$$\begin{aligned}\int_{-L}^{L}\int_{-L}^{L} \mathrm{d}x_1\mathrm{d}x_2 = &\int_{-L}^{0}\int_{-L}^{0} \mathrm{d}x_1\mathrm{d}x_2 + \int_{-L}^{0}\int_{0}^{L} \mathrm{d}x_1\mathrm{d}x_2 \\ &+ \int_{0}^{L}\int_{-L}^{0} \mathrm{d}x_1\mathrm{d}x_2 + \int_{0}^{L}\int_{0}^{L} \mathrm{d}x_1\mathrm{d}x_2 .\end{aligned} \tag{4.40}$$

The first term on the right corresponds to the situation in which both particles occupy the left half of the box.)

Thus, the probability $p^<$ that all the particles are found within the original volume V in the final state is given by

$$p^< = \frac{C_M^i}{C_M^f} = e^{(S_i - S_f)/k_B} . \tag{4.41}$$

This is an example of (2.20). For $N = 10^{24}$, $p^< = 2^{-10^{24}} \approx 10^{-0.3\times 10^{24}}$, an extremely small number. For all practical purposes, the process of expansion of the gas is irreversible. However, if $N = 4$, then $p^< = 2^{-4}$ and a spontaneous reversal of the process is not too unlikely. In this way, *the irreversible approach to an equilibrium state acquires a probabilistic interpretation*.

Exercise 4.1. Consider an isolated system consisting of two compartments of equal volume that are separated by a diathermal, rigid, impermeable partition held fixed in place. Initially, one of the compartments is filled with N noninteracting identical particles of species A, while the other is filled with N noninteracting identical particles of species B. If the partition is removed and a new state of equilibrium is reached, what is the increase in the total entropy? What would be the increase in the entropy if species A and B happen to be the same? ///

In the following two exercises, we find an explicit expression for $\mathscr{W}$ and $\overline{\Omega}$ for an ideal gas. Exercise 4.2 introduces a new mathematical tool, which will be used in Exercise 4.3.

Exercise 4.2. The **Gamma function** is defined by

$$\Gamma(s) := \int_0^\infty x^{s-1} e^{-x} \mathrm{d}x \quad (s > 0) . \tag{4.42}$$

Prove the following identities:

a.

$$\Gamma(s+1) = s\Gamma(s) . \tag{4.43}$$

b.

$$\Gamma(n+1) = n! , \tag{4.44}$$

where n is a positive integer.

c.

$$\Gamma\left(n + \frac{1}{2}\right) = \frac{\sqrt{\pi}(2n)!}{2^{2n} n!} . \tag{4.45}$$

///

Exercise 4.3. Following the steps indicated below, evaluate the partition function C_M for a system of N noninteracting identical particles confined to a box of volume V:

a. Calculate

$$I_n := \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-({x_1}^2+\cdots+{x_n}^2)} \mathrm{d}x_1 \cdots \mathrm{d}x_n \,. \tag{4.46}$$

b. Let U_n be the volume of the unit sphere in n-dimensional space. Then, the volume of the n-dimensional sphere of radius r is $U_n r^n$. Use this fact to express I_n in terms of U_n.

c. Show that U_n is given by

$$U_n = \frac{\pi^{n/2}}{\Gamma(n/2+1)} \,. \tag{4.47}$$

Check the validity of the formula for $n = 2$ and $n = 3$.

d. Show that

$$\mathscr{W}(E,V,N) = \frac{V^N}{N!} \left(\frac{2mE}{h^2} \right)^{3N/2} \int \theta \left(1 - \sum_{i=1}^{N} s_i^2 \right) \mathrm{d}s^N \,. \tag{4.48}$$

e. Evaluate C_M. ///

4.4 †Adiabatic Reversible Processes

Suppose that a system is subject to an external field, such as the one generated by a piston in Fig. 3.4, for example. The system is otherwise isolated. We know from thermodynamics that, if this external field varies very slowly, the process proceeds reversibly and the entropy S of the system remains constant. Our goal in this section is to provide a statistical mechanical demonstration of this fact.

As before, Hamiltonian of the system depends on the parameter λ denoting the position of the external body. We suppose that λ changes by an infinitesimal amount $\Delta\lambda$ over a long duration of time τ commencing at time t. Then, the total change in H can be expressed as

$$H(t+\tau) - H(t) = \int_t^{t+\tau} \frac{\mathrm{d}H}{\mathrm{d}t'} \mathrm{d}t' \,. \tag{4.49}$$

From (1.186), we have

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial t} \,. \tag{4.50}$$

For the process we are considering, the explicit time dependence of H arises solely from the time dependence of the parameter λ. Thus,

$$\frac{\partial H}{\partial t} = \frac{\partial H}{\partial \lambda} \frac{\mathrm{d}\lambda}{\mathrm{d}t} \,. \tag{4.51}$$

Using this expression in (4.49), we obtain

$$H(t+\tau) - H(t) = \int_t^{t+\tau} \frac{\partial H}{\partial \lambda} \frac{\mathrm{d}\lambda}{\mathrm{d}t'} \mathrm{d}t' \,. \tag{4.52}$$

If λ changes at a constant rate, $\mathrm{d}\lambda/\mathrm{d}t'$ may be written as $\Delta\lambda/\tau$, thus yielding

$$H(t+\tau) - H(t) = \frac{\Delta\lambda}{\tau}\int_t^{t+\tau}\frac{\partial H}{\partial\lambda}\mathrm{d}t' \,. \tag{4.53}$$

Because the change is supposed to occur very slowly, we let $\tau \to \infty$ and recall (3.4) to write[21]

$$H(t+\tau) - H(t) = \Delta\lambda\left\langle\frac{\partial H}{\partial\lambda}\right\rangle . \tag{4.54}$$

In a microcanonical ensemble, all copies in the ensemble have nearly identical energies. Thus, with very high degree of accuracy, we can write

$$H(t+\tau) = \langle H(t+\tau)\rangle = U(t+\tau) \quad \text{and} \quad H(t) = \langle H(t)\rangle = U(t) \,. \tag{4.55}$$

as in (4.21). Thus, (4.54) finally becomes

$$U(t+\tau) - U(t) = \mathrm{d}\lambda\left\langle\frac{\partial H}{\partial\lambda}\right\rangle , \tag{4.56}$$

where we replaced $\Delta\lambda$ by the usual notation $\mathrm{d}\lambda$ to indicate the infinitesimal change. Since the microcanonical ensemble is obtained as an approximation to canonical ensemble, there is nothing that prevents us from using (3.63), which in light of (4.56) indicates that $\mathrm{d}S = 0$.

The key assumption in arriving at this result was that the process occurs very slowly. But, $\tau \to \infty$ only means that τ is much larger compared to the characteristic time scale of molecular motions. Thus, the very slow process we considered here can be quite rapid when measured in a time scale of experiments. In fact, many textbooks on thermodynamics contend that $\mathrm{d}S \approx 0$ during a very *rapid* expansion or compression of a gas. There, the process is said to be rapid if no appreciable heat flow is involved between the system and the surroundings. These rapid processes are still very slow compared to the characteristic time scale of molecular motions.

4.5 Canonical Ensemble

We introduced microcanonical ensemble as an approximation to canonical ensemble. Historically, though, microcanonical ensemble was the first to be formulated. This is because an isolated system is the most natural system to consider in classical mechanics. From the point of view of practical applications, an isolated system is a rather rare occurrence. So, it is nice to have statistical ensembles suitable for describing systems in contact with their surroundings.

The situation is exactly the same as that in thermodynamics. There, the second law was formulated first for an isolated system. Then, various free energies were obtained by means of Legendre transformation of the energy representation.

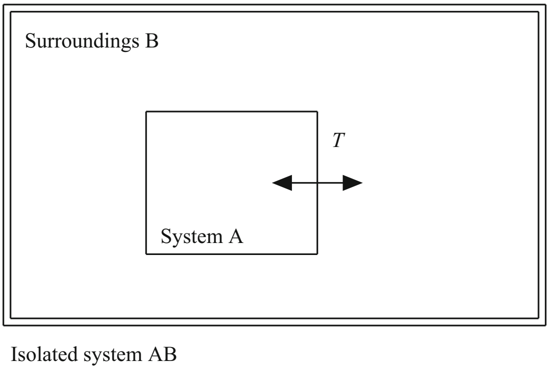


Fig. 4.2 The system of interest A is held at a constant temperature T by exchanging energy with the surroundings B. The composite system AB is isolated from the rest of the universe by means of an adiabatic, rigid, and impermeable wall.

Similarly, various statistical ensembles can be derived from the microcanonical ensemble. For this reason, the microcanonical ensemble is important conceptually despite the computational awkwardness you saw in Exercise 4.3.

4.5.1 Closed System Held at a Constant Temperature

Let us take a closed system A held at a constant temperature T. The number of particles N in A and the volume of A are both constant. But the system is allowed to exchange energy with the surroundings, which is assumed to be macroscopic. As we saw in Sect. 2.13.1, this thermal contact is what maintains the temperature of the system at a desired value.

Due to the energy exchange with the surroundings, the energy E_a of A fluctuates over time. Our immediate goal is to find the probability $p(E_a)\mathrm{d}E_a$ that the energy of A lies somewhere in the interval $(E_a - \mathrm{d}E_a, E_a]$. We use microcanonical ensemble for this purpose.

As we have indicated already, a microcanonical ensemble is suitable for describing an isolated system, which system A is not. In order to make the microcanonical ensemble applicable to the current situation, we construct a composite system AB consisting of A and the macroscopic surroundings B, and then isolate AB from the rest of the universe by means of an adiabatic, rigid, and impermeable wall as shown in Fig. 4.2. We shall assume that the interaction between A and B is sufficiently weak in the sense that

$$H_{ab} \approx H_a + H_b \,. \tag{4.57}$$

Because the composite system AB is isolated, the microcanonical ensemble applies. As in (4.16), we have

$$E_{ab} - \Delta E < H_{ab} \leq E_{ab} \tag{4.58}$$

for the allowed values of H_{ab}.

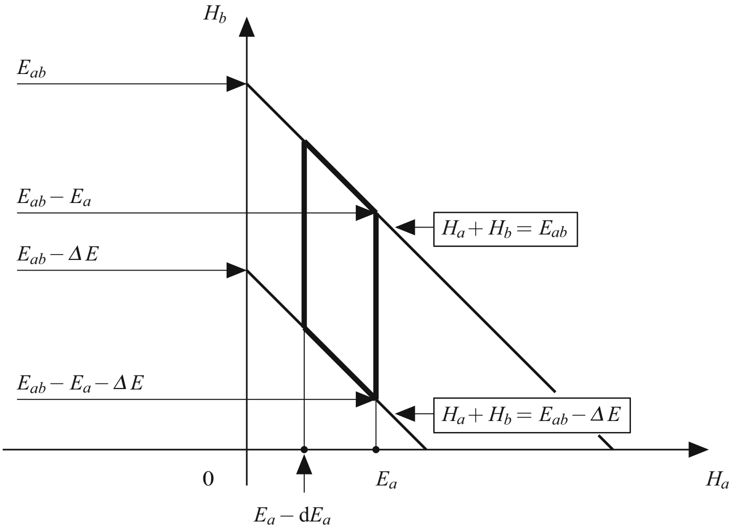


Fig. 4.3 The parallelogram region defined by four lines $H_a + H_b = E_{ab}$, $H_a + H_b = E_{ab} - \Delta E$, $H_a = E_a$, and $H_a = E_a - \mathrm{d}E_a$ contains all the microstates accessible to the composite system AB that are consistent with (4.58) and (4.59).

We recall that all the microstates of AB consistent with this condition on H_{ab} are equally probable. Thus, to compute $p(E_a)\mathrm{d}E_a$, we need to figure out the total number of microstates of AB that are consistent with both (4.58) *and* the condition

$$E_a - \mathrm{d}E_a < H_a \leq E_a \,. \tag{4.59}$$

The desired probability is this number divided by the total number of microstates of AB, which is simply $\overline{\Omega}_{ab}(E_{ab})\Delta E$.

Now, the number of microstates accessible to A under the condition (4.59) is simply

$$\overline{\Omega}_a(E_a)\mathrm{d}E_a \,, \tag{4.60}$$

where we suppressed the dependence of $\overline{\Omega}_a$ on the system volume and the number of mechanical degrees of freedom of the system as they are constant in this case. But, when A is at a particular microstate with a given energy, B can still be in various different microstates that are consistent with the remaining energy left to it.

As seen in Fig. 4.3, the conditions (4.58) and (4.59) specify the allowed values of H_b. The figure indicates that the interval for H_b depends on the exact value of H_a. This leads to a complication in figuring out the number of microstates accessible to system B. However, for sufficiently small $\mathrm{d}E_a$, we may replace the parallelogram indicated by thick solid lines by the rectangle defined by $H_a = E_a$, $H_a = E_a - \mathrm{d}E_a$, $H_b = E_{ab} - E_a$, and $H_b = E_{ab} - E_a - \Delta E$. Thus, we may write

$$E_{ab} - E_a - \Delta E < H_b \leq E_{ab} - E_a \,. \tag{4.61}$$

Accordingly, the number of microstates accessible to B that are consistent with both the conditions (4.58) and (4.59) is given by

$$\overline{\Omega}_b(E_{ab}-E_a)\Delta E\,. \tag{4.62}$$

The total number of microstates accessible to the composite system AB as a whole is then the product of (4.60) and (4.62):

$$\overline{\Omega}_a(E_a)\mathrm{d}E_a\overline{\Omega}_b(E_{ab}-E_a)\Delta E\,. \tag{4.63}$$

So, the desired probability is

$$p(E_a)\mathrm{d}E_a = \frac{\overline{\Omega}_b(E_{ab}-E_a)\Delta E\overline{\Omega}_a(E_a)\mathrm{d}E_a}{\overline{\Omega}_{ab}(E_{ab})\Delta E}\,. \tag{4.64}$$

The normalization condition of $p(E_a)$ leads to

$$\overline{\Omega}_{ab}(E_{ab})\Delta E = \int \overline{\Omega}_b(E_{ab}-E_a)\Delta E\overline{\Omega}_a(E_a)\mathrm{d}E_a\,, \tag{4.65}$$

where the integration is with respect to all possible values of E_a. This result is quite natural. In order to find out the total number of microstates accessible to AB, we first figure out the number of microstates accessible to AB when H_a is within a certain interval indicated by (4.59). This gives (4.63). Once the number of microstates is obtained for all the other intervals of width $\mathrm{d}E_a$, we can simply add up the results to obtain the total number of microstates.

4.5.2 Canonical Distribution

As it is, (4.64) is not very useful for practical computations. As we saw in Exercise 4.3, $\overline{\Omega}$ is a very awkward function to work with. To transform (4.64) a bit further, we recall Boltzmann's entropy formula (4.26) and write

$$\overline{\Omega}_b(E_{ab}-E_a)\Delta E = e^{\ln\overline{\Omega}_b(E_{ab}-E_a)\Delta E} = e^{S_b(E_{ab}-E_a)/k_B}\,. \tag{4.66}$$

Expanding $S_b(E_{ab}-E_a)$ in a Taylor series, we find

$$S_b(E_{ab}-E_a) = S_{b0} - \left.\frac{\partial S_b(E_b)}{\partial E_b}\right|_0 E_a + \text{h.o.} = S_{b0} - \frac{E_a}{T} + \text{h.o.}\,, \tag{4.67}$$

where the subscript 0 on S_b and its derivative indicates that they are to be computed under the condition that $E_b = E_{ab}$, while T defined in the last step is the temperature of the surroundings B when its energy is E_{ab}.

At this point, we note that the surroundings B can be chosen arbitrarily large compared to A, and hence the numbers of mechanical degrees of freedom of these two systems can be made to satisfy $f_a \ll f_b$. This implies that $E_{ab} \gg E_a$ for all reasonable values of E_a and hence T is quite insensitive to the exact value of E_a. This being the case, T can be referred to as the temperature of B regardless of the exact value of E_a and we can safely ignore the higher order terms in the Taylor series expansion of S_b.[22]

Recalling that $\beta := 1/k_BT$, we can rewrite (4.66) as

$$\overline{\Omega}_b(E_{ab}-E_a)\Delta E = e^{S_{b0}/k_B}e^{-\beta E_a} \ , \tag{4.68}$$

in terms of which (4.64) and (4.65) become

$$p(E_a)\mathrm{d}E_a = \frac{e^{S_{b0}/k_B}e^{-\beta E_a}\overline{\Omega}_a(E_a)\mathrm{d}E_a}{\overline{\Omega}_{ab}(E_{ab})\Delta E} \tag{4.69}$$

and

$$\overline{\Omega}_{ab}(E_{ab})\Delta E = e^{S_{b0}/k_B}\int \overline{\Omega}_a(E_a)e^{-\beta E_a}\mathrm{d}E_a \ , \tag{4.70}$$

respectively.

In what follows, we will be concerned only with system A. Thus, we drop the subscript a without a risk of confusion. Combining the last two equations,

$$p(E)\mathrm{d}E = \frac{1}{Z}\overline{\Omega}(E)e^{-\beta E}\mathrm{d}E \ , \tag{4.71}$$

in which

$$Z := \int \overline{\Omega}(E)e^{-\beta E}\mathrm{d}E \tag{4.72}$$

is the **canonical partition function** .[23] In Sect. 4.5.3, we see that (4.72) agrees with (3.162).

The free energy associated with this ensemble, by definition, is the **Helmholtz free energy**:

$$F = -k_BT\ln Z \ . \tag{4.73}$$

This definition, along with (2.172), leads to Gibbs's entropy formula as you will see in Exercise 4.5. Note that $\overline{\Omega}$ is a function of E, the volume V, and the number of mechanical degrees of freedom f of the system. When it is multiplied by the Boltzmann factor and integrated with respect to E, the final answer Z and hence F will be functions of T, V, and f. That is, (4.73) is a fundamental equation of the system.

We recall that $\overline{\Omega}(E)$ is the density of states. So, $\overline{\Omega}(E)\mathrm{d}E$ is the number of microstates of system A within the interval $(E-\mathrm{d}E, E]$. We already remarked that each of these microstates are equally probable.[24] Accordingly, (4.71) has the following interpretation. The factor $e^{-\beta E}/Z$ is the probability that the system A is at a *particular* microstate with energy somewhere in the interval $(E-\mathrm{d}E, E]$. But, there are $\overline{\Omega}(E)\mathrm{d}E$ equally probable microstates within this interval. Therefore, the probability that the energy of system A is in the interval $(E-\mathrm{d}E, E]$, irrespective of which particular microstate it is in, is given by (4.71).

Finally, our development up to this point is based entirely on classical mechanics, except for the $1/h^f \mathscr{P}$ factor originating from quantum mechanical considerations. Nevertheless, a fully quantum mechanical formulation also leads to (4.71) and (4.72). That is, they are valid regardless of the underlying mechanics. See Chap. 8 for details.

Example 4.2. Harmonic oscillator: In Exercise 3.8, you computed the constant volume heat capacity C_V of a diatomic molecule. The result was that, if the chemical bond connecting the two atoms is perfectly rigid, then, $C_V = \frac{5}{2}k_B$. The equipartition theorem indicates that, of this amount, $\frac{3}{2}k_B$ reflects the translational degrees of freedom of the center of mass of the molecule, while the remaining k_B comes from the rotation of the molecule around its center of mass. On the other hand, if the chemical bond is modeled as a harmonic spring, $C_V = \frac{7}{2}k_B$. The additional k_B comes from the vibrational motion (stretching and compressing) of the spring.

Experimentally, it is found that $C_V = \frac{5}{2}k_B$ for many (but not all) molecules at room temperature. Should we believe, then, that the chemical bond is absolutely rigid? So, if we are to pull on the atoms, then, the molecule will stay rigid until it suddenly snaps? This does not sound reasonable at all. Just as puzzling is the following observation: Physically, we expect that the harmonic spring will approach the rigid rod if the spring constant k is made infinitely large. However,

$$\lim_{k\to\infty} \frac{7}{2}k_B \neq \frac{5}{2}k_B \ . \tag{4.74}$$

This rather unsatisfactory state of affair stems from the fact that the chemical bond must be described *not* by classical mechanics but by quantum mechanics. (The same remark applies to both translational and rotational degrees of freedom of the molecule. At room temperature, however, the classical treatment proves sufficiently accurate. For details, see Chap. 8 of Ref. [1], for example.)

According to quantum mechanics, the energy E of the harmonic oscillator, a particle of mass m attached to a spring with spring constant k, can take only a discrete set of values as prescribed by the formula:

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) , \quad n = 0, 1, 2, \dots , \tag{4.75}$$

where

$$\omega := \sqrt{\frac{k}{m}} \quad \text{and} \quad \hbar := \frac{h}{2\pi} \ . \tag{4.76}$$

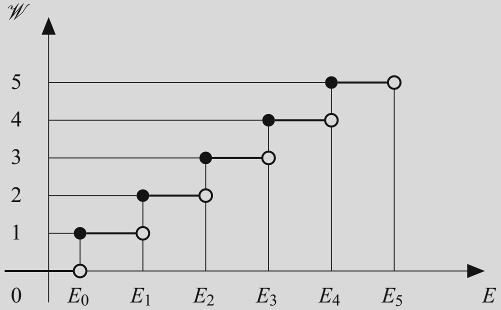
(See Sect. 8.14 for a derivation of this result.) In the case of a diatomic molecule consisting of atoms of mass m_1 and m_2, $m := m_1 m_2/(m_1 + m_2)$ is the reduced mass.

As remarked above, (4.72) is still applicable even for a quantum mechanical system. To compute Z, then, all we have to do is to figure out $\overline{\Omega}(E)$. It is actually easier to compute $\mathscr{W}(E)$, the number of states whose energy is *less than or equal to* E, and then obtain $\overline{\Omega}(E)$ as the partial derivative of $\mathscr{W}(E)$ with respect to E.

To find $\mathscr{W}(E)$, we proceed as follows. If E is less than E_0, there is no quantum mechanical state. Thus $\mathscr{W}(E) = 0$ if $E < E_0$. But, there is a single quantum state at $E = E_0$ and then another at $E = E_1$. Thus, $\mathscr{W}(E) = 1$ if $E_0 \leq E < E_1$. Likewise, $\mathscr{W}(E) = 2$ if $E_1 \leq E < E_2$. In this way, we end up with

$$\mathscr{W}(E) = \sum_{n=0}^{\infty} \theta(E - E_n) . \tag{4.77}$$

The graph of $\mathscr{W}(E)$ looks like this:



Taking the derivative with respect to E, we find

$$\overline{\Omega}(E) = \frac{\partial \mathscr{W}(E)}{\partial E} = \sum_{n=0}^{\infty} \delta(E - E_n) . \tag{4.78}$$

Substituting this expression in (4.72),

$$Z = \int_0^{\infty} \sum_{n=0}^{\infty} \delta(E - E_n) e^{-\beta E} \mathrm{d}E = \sum_{n=0}^{\infty} \int_0^{\infty} \delta(E - E_n) e^{-\beta E} \mathrm{d}E , \tag{4.79}$$

where the lower limit of the integration can be anything as long as it is less than E_0 since $\overline{\Omega}(E) \equiv 0$ for $E < E_0$. Using (D.15), we find

$$Z = \sum_{n=0}^{\infty} e^{-\beta E_n} = \sum_{n=0}^{\infty} e^{-\beta\hbar\omega\left(n+\frac{1}{2}\right)} = e^{-\beta\hbar\omega/2} \sum_{n=0}^{\infty} \left(e^{-\beta\hbar\omega}\right)^n . \tag{4.80}$$

To carry out the indicated summation, we recall (B.25). Setting $r = e^{-\beta\hbar\omega}$, which is less than unity since $\beta\hbar\omega > 0$, we find

$$Z = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}} . \tag{4.81}$$

Thus,

$$\ln Z = -\frac{\beta\hbar\omega}{2} - \ln\left(1 - e^{-\beta\hbar\omega}\right) \tag{4.82}$$

and

$$U = -\frac{\partial \ln Z}{\partial \beta} = \frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} . \tag{4.83}$$

So, the heat capacity of the quantum harmonic oscillator is given by

$$C_{\text{ho}} = \frac{\partial U}{\partial T} = \frac{\mathrm{d}\beta}{\mathrm{d}T}\frac{\partial U}{\partial \beta} = -\frac{1}{k_B T^2}\frac{-(\hbar\omega)^2 e^{\beta\hbar\omega}}{\left(e^{\beta\hbar\omega} - 1\right)^2} = k_B(\beta\hbar\omega)^2\frac{e^{\beta\hbar\omega}}{\left(e^{\beta\hbar\omega} - 1\right)^2} . \tag{4.84}$$

Let us define Θ_v by

$$\beta\hbar\omega = \frac{1}{T}\frac{\hbar}{k_B}\sqrt{\frac{k}{m}} =: \frac{\Theta_v}{T} \tag{4.85}$$

and write

$$C_{\text{ho}} = k_B\left(\frac{\Theta_v}{T}\right)^2\frac{e^{\Theta_v/T}}{\left(e^{\Theta_v/T} - 1\right)^2} . \tag{4.86}$$

Suppose that $\Theta_v/T \ll 1$, which occurs if T is large, m is large, or k is small. Then,

$$e^{\Theta_v/T} \approx 1 + \frac{\Theta_v}{T} + \frac{1}{2}\left(\frac{\Theta_v}{T}\right)^2 , \tag{4.87}$$

and hence to the first order of Θ_v/T,

$$C_{\text{ho}} \approx k_B\left(\frac{\Theta_v}{T}\right)^2\frac{1 + \Theta_v/T}{\left[1 + \Theta_v/T + 1/2(\Theta_v/T)^2 - 1\right]^2} = k_B \tag{4.88}$$

in agreement with the classical result. With some extra work, we can find that

$$C_{\text{ho}} \approx k_B\left[1 - \frac{1}{12}\left(\frac{\Theta_v}{T}\right)^2\right] . \tag{4.89}$$

If, on the other hand, $\Theta_v/T \gg 1$, then $e^{\Theta_v/T} \gg 1$, and we have

$$C_{\text{ho}} \approx k_B\left(\frac{\Theta_v}{T}\right)^2\frac{e^{\Theta_v/T}}{\left(e^{\Theta_v/T}\right)^2} = k_B\left(\frac{\Theta_v}{T}\right)^2 e^{-\Theta_v/T} , \tag{4.90}$$

which becomes vanishingly small in the low T limit. Temperature dependence of C_{ho} is illustrated in Fig. 4.4.

We note that Θ_v is a material property. Some representative values, along with the value of C_{ho}/k_B computed from (4.86), are given in Table 4.1.

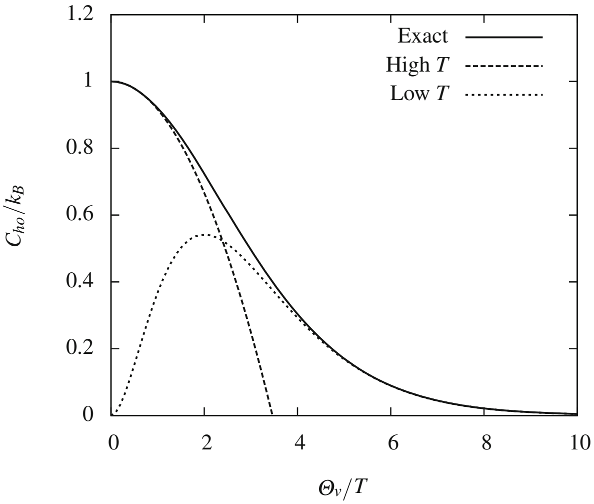


Fig. 4.4 Temperature dependence of the heat capacity C_{ho} of a quantum harmonic oscillator. "Exact, " "High T," and "Low T" refer to (4.86), (4.89), and (4.90), respectively. A classical treatment of the harmonic oscillator gives $C_{\mathrm{ho}}/k_B = 1$ as you saw in Exercise 3.8c.

Thus, except for Cl_2 and I_2, the vibrational motion makes a negligible contribution to C_V at room temperature. This is why the rigid chemical bond model works so well in predicting C_V for many (but not all) diatomic molecules. At the room temperature, the vibrational degrees of freedom of the chemical bond is still frozen.

Quantum mechanical treatment of the rotational degrees of freedom of a symmetric diatomic molecule, such as H_2 in the table, requires a special consideration. See Ref. [1] for details.

Exercise 4.4. Prove (4.89). ///

	Θ_v (K)	C_{ho}/k_B at 300 K
H_2	6215	4.32×10^{-7}
HCl	4227	1.51×10^{-4}
N_2	3374	1.65×10^{-3}
CO	3100	3.47×10^{-3}
Cl_2	810	0.563
I_2	310	0.916

Table 4.1 Representative examples of Θ_v values and C_{ho} at 300 K. The entries for Θ_v are adapted from *Statistical Mechanics* by Donald McQuarrie, and are reprinted with permission from University Science Books, all rights reserved [2].

4.5.3 Classical Canonical Partition Function

The quantity Z defined by (4.72) is actually the canonical partition function introduced earlier. Using the expression for $\overline{\Omega}(E)$ given by (4.32) in (4.72):

$$\begin{aligned}
Z &= \int \left[\frac{1}{h^f \mathscr{P}} \int \delta\left(E - H(q^f, p^f)\right) \mathrm{d}q^f \mathrm{d}p^f \right] e^{-\beta E} \mathrm{d}E \\
&= \frac{1}{h^f \mathscr{P}} \int \left[\int \delta\left(E - H(q^f, p^f)\right) e^{-\beta E} \mathrm{d}E \right] \mathrm{d}q^f \mathrm{d}p^f \\
&= \frac{1}{h^f \mathscr{P}} \int e^{-\beta H(q^f, p^f)} \mathrm{d}q^f \mathrm{d}p^f \ ,
\end{aligned} \tag{4.91}$$

which is (3.161). Since Z is the normalization constant and $\int$ represents the sum over states, this result implies that the probability of finding a system within an infinitesimal volume element $\mathrm{d}q^f \mathrm{d}p^f$ taken around (q^f, p^f) is given by

$$\rho(q^f, p^f) \mathrm{d}q^f \mathrm{d}p^f = \frac{1}{Z} e^{-\beta H(q^f, p^f)} \frac{\mathrm{d}q^f \mathrm{d}p^f}{h^f \mathscr{P}} \ , \tag{4.92}$$

which is (3.165). Unlike (4.71) and (4.72), the validity of (4.91) and (4.92) is limited to classical mechanical systems only.

Exercise 4.5. We now have two definitions for the Helmholtz free energy, one is (2.172) from thermodynamics and the other is (4.73). Consolidate them using (4.71). A similar attempt for (4.92) lead to (3.168) as you saw in Exercise 3.14. ///

4.5.4 Applicability of Canonical Ensemble

Looking back at our derivation, we can list two important assumptions made in arriving at the canonical ensemble. First, we assumed that the interaction between A and B is sufficiently weak. As discussed in Sect. 3.11, an interaction is said to be sufficiently weak if it allows exchange of energy between A and B *and* if the Hamiltonian of the composite system AB can be expressed as in (4.57). Second, we assumed that the number of the mechanical degrees of freedom of the surroundings f_b is much larger than that of the system of interest f_a. These are the only assumptions we made. Nowhere in our derivation have we assumed that f_a is large. Provided that these two conditions are met, the canonical ensemble is generally applicable even if $f_a = 1$.

Finally, let us remark on the statistical mechanical meaning of T. Because of exchange of energy between subsystem A and the surroundings B, the energy of system A fluctuates with time. Without the detailed knowledge of this interaction, it is not possible to predict how E_a evolves with time. Nevertheless, if the behavior of A is observed for a long interval of time, the resulting distribution of the energy

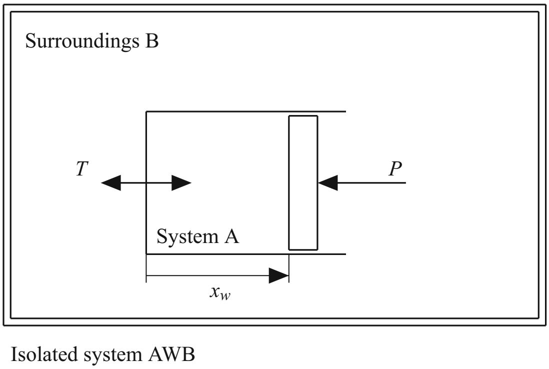


Fig. 4.5 The system of interest A is held at constant temperature T and pressure P due to exchange of energy and repartitioning of volume with the surroundings B. The composite system AWB is isolated from the rest of the universe by means of an adiabatic, rigid, and impermeable wall.

$p(E_a)$ is given by (4.71). The temperature T appears here simply as a *parameter* characterizing this distribution.

4.6 ‡Isothermal–Isobaric Ensemble

In this section, we derive an appropriate ensemble for describing a closed system held at a given temperature T and a pressure P. Our system A is the region inside a cylinder fitted with a piston. The quantity of interest is the probability $p(E_a,V_a)\mathrm{d}E_a\mathrm{d}V_a$ that A has the energy in $(E_a-\mathrm{d}E_a,E_a]$ and the volume in $(V_a,V_a+\mathrm{d}V_a]$. As with our derivation of canonical ensemble, we use a microcanonical ensemble as our starting point. Thus, we consider the composite system AWB consisting of system A, the piston W, and the surrounding B, and then isolate AWB from the rest of the universe by means of an adiabatic, rigid, and impermeable wall. See Fig. 4.5. Under the usual assumption of weak interaction, the Hamiltonian H_{awb} of the composite system may be written as

$$H_{awb} = H'_a + H'_b + K_w + \phi_{aw} + \phi_{bw} \ , \tag{4.93}$$

where we use $'$ to indicate that H'_a contains the generalized coordinates and their conjugate momenta pertaining only to the degrees of freedom of system A but not to the degrees of freedom of the piston. Likewise for H'_b. $K_w := p_w^2/2m_w$ is the kinetic energy of the piston, ϕ_{aw} denotes the potential energy due to the interaction between system A and the piston, and likewise for ϕ_{bw}. Now, we set $H_a := H'_a + \phi_{aw}$ and $H_b := H'_b + \phi_{bw}$ and write

$$H_{awb} = H_a + H_b + K_w \ . \tag{4.94}$$

Applying (4.32) to the composite system AWB, we see that the total number of microstates accessible to the composite system as a whole is given by

$$\overline{\Omega}_{awb}(E_{awb})\Delta E = \int \delta(E_{awb} - H_{awb})\Delta E \frac{\mathrm{d}q^{f_a}\mathrm{d}p^{f_a}}{h^{f_a}\mathscr{P}_a}\frac{\mathrm{d}x_w\mathrm{d}p_w}{h}\frac{\mathrm{d}q^{f_b}\mathrm{d}p^{f_b}}{h^{f_b}\mathscr{P}_b}\,, \tag{4.95}$$

where x_w and p_w refer to the position and the linear momentum of the piston, respectively. Using (D.26),

$$\begin{aligned}\delta(E_{awb} - H_{awb}) &= \delta(E_{awb} - H_a - H_b - K_w)\\ &= \int \delta(E_{awb} - E_a - K_w - H_b)\delta(E_a - H_a)\mathrm{d}E_a\,.\end{aligned} \tag{4.96}$$

Substituting this expression into (4.95), we identify

$$\overline{\Omega}_a(E_a, x_w) = \int \delta(E_a - H_a)\frac{\mathrm{d}q^{f_a}\mathrm{d}p^{f_a}}{h^{f_a}\mathscr{P}_a} \tag{4.97}$$

as the density of states for A. The indicated x_w dependence arises through that of H_a. Similarly,

$$\overline{\Omega}_b(E_{awb} - E_a - K_w, x_w) = \int \delta(E_{awb} - E_a - K_w - H_b)\frac{\mathrm{d}q^{f_b}\mathrm{d}p^{f_b}}{h^{f_b}\mathscr{P}_b}\,. \tag{4.98}$$

It follows that

$$\overline{\Omega}_{awb}(E_{awb})\Delta E = \iiint \overline{\Omega}_b(E_{awb} - E_a - K_w, x_w)\Delta E\,\overline{\Omega}_a(E_a, x_w)\mathrm{d}E_a\frac{\mathrm{d}x_w\mathrm{d}p_w}{h}\,. \tag{4.99}$$

In this equation, $\overline{\Omega}_b\Delta E$ is the number of microstates for system B when its energy is in the interval $(E_{awb} - E_a - K_w - \Delta E, E_{awb} - E_a - K_w]$ and piston W is at x_w. Similarly, $\overline{\Omega}_a\mathrm{d}E_a$ is the number of microstates for system A when its energy is in $(E_a - \mathrm{d}E_a, E_a]$ with the piston at x_w. According to (4.99), $\mathrm{d}x_w\mathrm{d}p_w/h$ is then the number of microstates for the piston when its phase point is somewhere in the infinitesimal phase volume $\mathrm{d}x_w\mathrm{d}p_w$. This result is consistent with the interpretation we have given to the quantum mechanical correction factor due to the Heisenberg uncertainty principle.

As before, we use Boltzmann's entropy formula

$$\overline{\Omega}_b(E_{awb} - E_a - K_w, x_w)\Delta E = e^{S_b(E_{awb} - E_a - K_w, x_w)/k_B} \tag{4.100}$$

and expand S_b in Taylor series:

$$S_b(E_{awb} - E_a - K_w, x_w) = S_{b0} - \left.\frac{\partial S_b}{\partial E_b}\right|_0 (E_a + K_w) + \left.\frac{\partial S_b}{\partial x_w}\right|_0 x_w + \text{h.o.}, \tag{4.101}$$

where the subscript 0 on S_b and its partial derivatives indicates that they are to be computed at $E_b = E_{awb}$ and $x_w = 0$, that is, system A has zero volume and zero energy.

That such a situation is extremely unlikely to occur for nonempty system A is of no concern for us here. The indicated state servers only as a reference point around which to perform the Taylor series expansion of S_b, the entropy of system B. All we need is that S_b is well defined and differentiable at that state.

The first partial derivative in this equation is just $1/T$. In order to transform the second partial derivative into something familiar, let $A_w := V_a/x_w$ denote the cross sectional area of the cylinder. Then,

$$\left.\frac{\partial S_b}{\partial x_w}\right|_0 x_w = \left.\frac{\partial S_b}{\partial (A_w x_w)}\right|_0 A_w x_w = \left.\frac{\partial S_b}{\partial V_a}\right|_0 V_a = -\left.\frac{\partial S_b}{\partial V_b}\right|_0 V_a = -\frac{P}{T} V_a \,. \tag{4.102}$$

Thus, ignoring the higher order terms, we have

$$S_b(E_{awb} - E_a - K_w, x_w) = S_{b0} - \frac{E_a + K_w}{T} - \frac{PV_a}{T} \,. \tag{4.103}$$

Combining the results, we rewrite (4.99) as

$$\overline{\Omega}_{awb}(E_{awb})\Delta E = \iiint e^{S_{b0}/k_B} e^{-\beta E_a - \beta K_w - \beta P V_a} \overline{\Omega}_a(E_a, V_a) \mathrm{d}E_a \frac{\mathrm{d}x_w \mathrm{d}p_w}{h} \,. \tag{4.104}$$

Performing the integration with respect to p_w and replacing $\mathrm{d}x_w$ by $\mathrm{d}V_a/A_w$, we obtain

$$\overline{\Omega}_{awb}(E_{awb})\Delta E = e^{S_{b0}/k_B} \int e^{-\beta P V_a} \int \overline{\Omega}_a(E_a, V_a) e^{-\beta E_a} \frac{\mathrm{d}E_a \mathrm{d}V_a}{A_w \Lambda_w} \,, \tag{4.105}$$

where $\Lambda_w := h/\sqrt{2\pi m_w k_B T}$ is the thermal wavelength of the piston W. Because the $1/\Lambda_w$ factor arose upon the integration of $e^{-\beta K_w} \mathrm{d}x_w \mathrm{d}p_w/h$ with respect to p_w, $\mathrm{d}x_w/\Lambda_w = \mathrm{d}V_a/A_w\Lambda_w$ may be conveniently interpreted as the number of microstates accessible to the piston W when it is found somewhere between x_w and $x_w + \mathrm{d}x_w$ regardless of its momentum.

By carrying out the integrations with respect to E_a and V_a only over intervals of infinitesimal widths $\mathrm{d}E_a$ and $\mathrm{d}V_a$, we find the number of microstates accessible to the composite system AWB when H_a and V_A are within these infinitesimal intervals:

$$e^{S_{b0}/k_B} e^{-\beta P V_a} \overline{\Omega}_a(E_a, V_a) e^{-\beta E_a} \frac{\mathrm{d}E_a \mathrm{d}V_a}{A_w \Lambda_w} \,. \tag{4.106}$$

The desired probability, $p(E_a, V_a)\mathrm{d}E_a \mathrm{d}V_a$, is obtained by dividing this expression by (4.105).

Once again, we can drop the subscript a without a risk of confusion. Thus,

$$p(E,V)\mathrm{d}E\mathrm{d}V = \frac{1}{Y} e^{-\beta P V} \overline{\Omega}(E,V) e^{-\beta E} \frac{\mathrm{d}E \mathrm{d}V}{A_w \Lambda_w} \,, \tag{4.107}$$

where we defined the **isothermal–isobaric partition function** by

$$Y := \int e^{-\beta PV} \int \overline{\Omega}(E,V) e^{-\beta E} \frac{\mathrm{d}E\mathrm{d}V}{A_w \Lambda_w} . \tag{4.108}$$

The free energy associated with this ensemble is, by definition, the **Gibbs free energy**:

$$G := -k_B T \ln Y . \tag{4.109}$$

Recall that $\overline{\Omega}$ is a function of E, V, and the number of mechanical degrees of freedom f. When this is multiplied by the Boltzmann factor and $e^{-\beta PV}$ and then integrated with respect to E and V, we obtain a function of T, P, and f. The Gibbs free energy G also depends on this same set of variables. Equation (4.109) is therefore a fundamental equation of the system.

In (4.108), the integration with respect to V extends over all possible values. Because the surrounding can be made arbitrarily large, the upper limit of the integration over V is usually set to ∞. This is acceptable if the integrand vanishes sufficiently fast, which is usually the case if $P > 0$. On the other hand, the lower limit is set to zero, the theoretical minimum. Except for an ideal gas, $\overline{\Omega}(E,V)$ becomes negligibly small with increasing N/V unless E is extremely large. For example, imagine a densely packed liquid phase, in which there is little room for a molecule to move around. But, if E is very large, the Boltzmann factor effectively kills off the integrand and the integration with respect to V from $V = 0$ to ∞ is finite.

The integral

$$\int \overline{\Omega}(E,V) e^{-\beta E} \mathrm{d}E \tag{4.110}$$

in (4.108) is the canonical partition function Z of system A when its volume is V. The partition function Y is the Laplace transform of Z.

Using (4.32) in (4.108), we arrive at

$$Y = \frac{1}{h^f \mathscr{P}} \int_0^\infty e^{-\beta PV} \int e^{-\beta H(q^f,p^f)} \mathrm{d}q^f \mathrm{d}p^f \frac{\mathrm{d}V}{A_w \Lambda_w} . \tag{4.111}$$

Interpreting the integral signs as the sum over microstates, we identify

$$\rho(q^f,p^f,V)\mathrm{d}q^f \mathrm{d}p^f \mathrm{d}V = \frac{1}{Y} e^{-\beta PV} e^{-\beta H(q^f,p^f)} \frac{\mathrm{d}q^f \mathrm{d}p^f}{h^f \mathscr{P}} \frac{\mathrm{d}V}{A_w \Lambda_w} \tag{4.112}$$

as the probability that the piston W defines volume between V and $V + \mathrm{d}V$ and system A has its phase point somewhere in the phase volume $\mathrm{d}q^f \mathrm{d}p^f$ centered around (q^f, p^f).

Exercise 4.6. Deduce Gibbs entropy formula by combining (2.196), (4.107), and (4.109). Do the same using (4.112). ///

Because of the $1/A_w \Lambda_w$ factor, the isothermal-isobaric partition function is dimensionless. It is customary, however, to avoid an explicit reference to the piston, and redefine the partition function Y by

$$Y := \int e^{-\beta PV} \int \overline{\Omega}(E,V)e^{-\beta E}\mathrm{d}E\mathrm{d}V \ . \tag{4.113}$$

Using (4.32), we arrive at

$$Y = \frac{1}{h^f \mathscr{P}} \int_0^\infty e^{-\beta PV} \int e^{-\beta H(q^f,p^f)}\mathrm{d}q^f\mathrm{d}p^f\mathrm{d}V \ . \tag{4.114}$$

This partition function has the dimension of volume. When the Gibbs free energy is calculated as $-k_BT\ln Y$ using this dimensional Y, its numerical value depends on the unit used to measure length. A change in the unit of length amounts to an addition of a constant to the Gibbs free energy. This is of no consequence when discussing the difference in the Gibbs free energy between two states. We also note that $p(E,V)\mathrm{d}E\mathrm{d}V$ given by (4.107) remains unaffected when we drop the $1/A_w\Lambda_w$ factor. We have kept the factor up to this point since such a notion as the number of microstates becomes very difficult to conceive of otherwise.

Exercise 4.7. Find the partition functions Y (without $A_w\Lambda_w$) for an ideal gas, modeled as a system of N noninteracting identical particles. You need to recall the definition and properties of the Gamma function from Exercise 4.2. ///

Example 4.3. A system of three rods:[25] As shown in Fig. 4.6, a chain of three rod-shape molecules is confined to a cylinder of cross sectional area A fitted with a piston and is held under a constant temperature T and pressure P. The sequence of molecules ABC cannot be changed, but each molecule can take two distinct orientations, that is, horizontal (A and C in Fig. 4.6) and vertical (B in Fig. 4.6). The length of the molecule is l and the width is w. System energy E is always 0 independent of the orientations of the molecules. Let us find the equation of state, that is, an equation relating the average length $\langle L\rangle$ of the system to T and P.

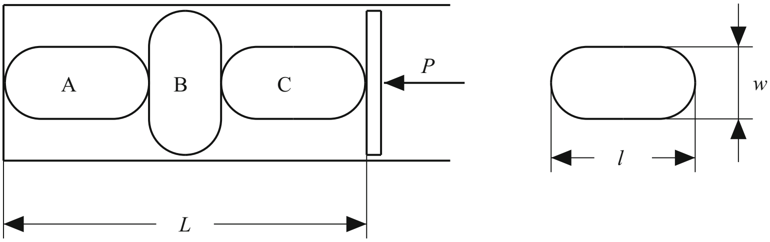


Fig. 4.6 A system of three rods held at a constant temperature and pressure.

The following orientations of molecules (A,B,C) are possible:

$$
\begin{array}{ll}
(w,w,w) & L=3w\,, \\
(l,w,w)\,,\ (w,l,w)\,,\ (w,w,l) & L=2w+l\,, \\
(l,l,w)\,,\ (l,w,l)\,,\ (w,l,l) & L=w+2l\,, \\
(l,l,l) & L=3l\,,
\end{array}
$$

where we also included the corresponding values of L. The energy is zero regardless of the configuration. Thus, the density of states is given by

$$
\begin{aligned}
&\overline{\Omega}(E,L) \\
&=\delta(E)\left[\delta(L-3w)+3\delta(L-2w-l)+3\delta(L-w-2l)+\delta(L-3l)\right].
\end{aligned}
\tag{4.115}
$$

To see this, we note that $\overline{\Omega}(E,L)$, upon integration over some domain on the EL-plane, gives the number of states within that domain. Thus, the point $(E,L)=(0,3w)$, if included, should contribute 1 to the integral while the point $(E,L)=(0,2w+l)$ should yield 3 upon integration. Equation (4.115) is compatible with this requirement.

Using (4.115) in (4.113),

$$
\begin{aligned}
Y&=\int_0^\infty e^{-\beta PV}\int\overline{\Omega}(E,L)e^{-\beta E}\mathrm{d}E\mathrm{d}V \\
&=\int_0^\infty e^{-\beta PAL}\left[\delta(L-3w)+3\delta(L-2w-l)\right. \\
&\qquad\qquad\left.+3\delta(L-w-2l)+\delta(L-3l)\right]A\mathrm{d}L \\
&=A(x^3+3x^2y+3xy^2+y^3)=A(x+y)^3\,,
\end{aligned}
\tag{4.116}
$$

where $x:=e^{-\beta PAw}$ and $y:=e^{-\beta PAl}$. Finally,

$$
\langle V\rangle=-\left(\frac{\partial\ln Y}{\partial\beta P}\right)_\beta=-\frac{3}{x+y}\left[\left(\frac{\partial x}{\partial\beta P}\right)_\beta+\left(\frac{\partial y}{\partial\beta P}\right)_\beta\right]=\frac{3A(wx+ly)}{x+y}\,.
\tag{4.117}
$$

Dividing by A, we find

$$
\langle L\rangle=\frac{3(wx+ly)}{x+y}\,,
\tag{4.118}
$$

which is the desired equation of state.

If you prefer, you can stick to the formalism we used in Example 4.2 and work with $\mathscr{W}(E,L)$, which is the number of states with the energy less than or equal to E *and* the length less than or equal to L. Note that no states are allowed unless $E=0$. So, $\mathscr{W}$ is zero if $E<0$ and positive if $E\geq 0$. This implies that $\mathscr{W}$ is proportional to $\theta(E)$.

Let $w < l$ and assume that $E = 0$. If $L < 3w$, $\mathscr{W} = 0$, while $\mathscr{W} = 1$ if $3w \leq L < 2w + l$. Proceeding similarly with larger values of L, we get

$$\begin{aligned}\mathscr{W}(E,L) = \theta(E)\,[&\theta(L-3w) + 3\theta(L-2w-l)\\ &+3\theta(L-w-2l) + \theta(L-3l)]\ . \end{aligned} \tag{4.119}$$

Note that the product $\theta(E)\theta(L-3w)$ is nonzero only if $E \geq 0$ and $L \geq 3w$. A similar remark applies to other products in (4.119). Equation (4.115) follows from

$$\overline{\Omega}(E,L) = \frac{\partial^2 \mathscr{W}}{\partial E \partial L}\ . \tag{4.120}$$

To see that this is a proper way to compute $\overline{\Omega}(E,L)$, note that

$$\begin{aligned}\int_{E_1}^{E_2}\int_{L_1}^{L_2}\overline{\Omega}(E,L)\mathrm{d}L\mathrm{d}E &= \int_{E_1}^{E_2}\left[\frac{\partial\mathscr{W}(E,L_2)}{\partial E} - \frac{\partial\mathscr{W}(E,L_1)}{\partial E}\right]\mathrm{d}E\\ &= \mathscr{W}(E_2,L_2) - \mathscr{W}(E_1,L_2) - \mathscr{W}(E_2,L_1) + \mathscr{W}(E_1,L_1)\ ,\end{aligned} \tag{4.121}$$

which may be recognized as the number of states with $E_1 < E \leq E_2$ *and* $L_1 < L \leq L_2$. You can convince yourself of this by drawing rectangles on the EL-plane.

Exercise 4.8. Generalize Example 4.3 for a system of N rods. ///

4.7 Grand Canonical Ensemble

In this section, we shall consider an open system which can exchange energy and particles of a given species with the surroundings. By means of this exchange, the temperature and the chemical potential of that species in the system are held at constant values.

Our formulation generalizes quite straightforwardly to a system open to multiple species, but the notation becomes quite unmanageable. Thus, we refrain from exploring this more general situation here. The formulation we present does *not* exclude mixtures as long as the other species are confined to the system of our interest.

As before, the quantity of interest is the probability $p(E_a,N_a)\mathrm{d}E_a$ that system A contains N_a particles and has the energy in $E_a - \mathrm{d}E_a < H_a \leq E_a$. The expression for the probability will be found using microcanonical ensemble. Once again, we construct a composite system AB consisting of the system of interest A and the surroundings B, and then isolate AB from the rest of the universe by means of an adiabatic, rigid, and impermeable wall as shown in Fig. 4.7.

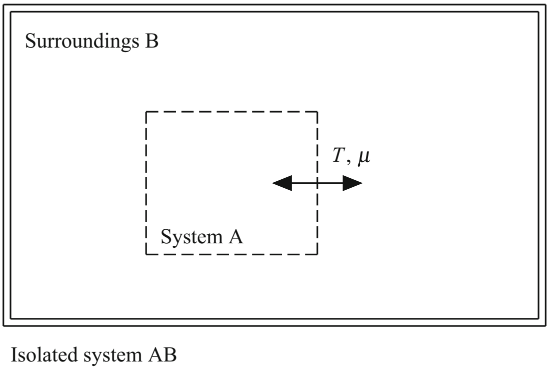


Fig. 4.7 The system of interest A is held at a constant temperature T and chemical potential μ due to exchange of energy and particles with the surroundings B. The composite system AB is isolated from the rest of the universe by means of an adiabatic, rigid, and impermeable wall.

The argument that led to (4.64) generalizes directly to the current situation simply by noting that $\overline{\Omega}_a$, $\overline{\Omega}_b$, and hence p now depends on N_a as well as on E_a:

$$p(E_a,N_a)\mathrm{d}E_a = \frac{\overline{\Omega}_b(E_{ab}-E_a,N_{ab}-N_a)\Delta E\,\overline{\Omega}_a(E_a,N_a)\mathrm{d}E_a}{\overline{\Omega}_{ab}(E_{ab},N_{ab})\Delta E}\,. \tag{4.122}$$

The normalization condition of $p(E_a,N_a)$ leads to

$$\overline{\Omega}_{ab}(E_{ab},N_{ab})\Delta E = \sum_{N_a}\int \overline{\Omega}_b(E_{ab}-E_a,N_{ab}-N_a)\Delta E\,\overline{\Omega}_a(E_a,N_a)\mathrm{d}E_a\,, \tag{4.123}$$

where the integral is taken over all possible values of E_a and the summation is over all possible values of N_a. The bounds on N_a will be specified later. We can interpret (4.123) by following the argument similar to the one given to (4.65).

In order to transform (4.122) to a more convenient form, we define the chemical potential μ of a particle in the surroundings by

$$-\frac{\mu}{T} := \left.\frac{\partial S_b(E_b)}{\partial N_b}\right|_0\,, \tag{4.124}$$

where we introduced the subscript 0 to indicate that quantities bearing them are to be computed at $E_b = E_{ab}$ and $N_b = N_{ab}$, that is, the surroundings have all the energy and particles of the composite system.

If the number of mechanical degrees of freedom of B is much larger than that of A, then the Taylor series expansion of $S_b(E_{ab}-E_a,N_{ab}-N_a)$ will be sufficiently accurate when only the first-order terms in E_a and N_a are retained. This gives

$$\overline{\Omega}_b(E_{ab}-E_a,N_{ab}-N_a)\Delta E = e^{S_{b0}/k_B}e^{-\beta E_a+\beta\mu N_a}\,, \tag{4.125}$$

in terms of which we rewrite (4.123) as

$$\overline{\Omega}_{ab}(E_{ab},N_{ab})\Delta E = e^{S_{b0}/k_B}\sum_{N_a} e^{\beta\mu N_a}\int \overline{\Omega}_a(E_a,N_a)e^{-\beta E_a}\mathrm{d}E_a\ . \tag{4.126}$$

As with other ensembles, we drop the subscript a here and substitute the last two equations into (4.122):

$$p(E,N)\mathrm{d}E = \frac{1}{\Xi}e^{\beta\mu N}\overline{\Omega}(E,N)e^{-\beta E}\mathrm{d}E\ , \tag{4.127}$$

where

$$\Xi = \sum_N e^{\beta\mu N}\int \overline{\Omega}(E,N)e^{-\beta E}\mathrm{d}E \tag{4.128}$$

is the **grand canonical partition function**. The free energy associated with this ensemble, by definition, is the **grand potential**:[26]

$$\Omega := -k_BT\ln\Xi\ . \tag{4.129}$$

Recall that $\overline{\Omega}$ is a function of E, V, N, and the number f of any remaining mechanical degrees of freedom. Any other species, if present in the system, afford an example of such remaining degrees of freedom. Thus, Ξ and Ω are functions of T, V, μ, and f. Equation (4.129) is then a fundamental equation of the system.

For a classical mechanical system of identical particles, (4.128) may be written as

$$\Xi = \sum_N \frac{e^{\beta\mu N}}{h^{3N}N!}\int e^{-\beta H(\boldsymbol{r}^N,\boldsymbol{p}^N)}\mathrm{d}\boldsymbol{r}^N\mathrm{d}\boldsymbol{p}^N\ . \tag{4.130}$$

We interpret $\sum_N\int$ as the sum over states. Then,

$$\rho(\boldsymbol{r}^N,\boldsymbol{p}^N,N)\mathrm{d}\boldsymbol{r}^N\mathrm{d}\boldsymbol{p}^N = \frac{1}{\Xi}e^{\beta\mu N}e^{-\beta H(\boldsymbol{r}^N,\boldsymbol{p}^N)}\frac{\mathrm{d}\boldsymbol{r}^N\mathrm{d}\boldsymbol{p}^N}{h^{3N}N!} \tag{4.131}$$

is the probability that system A contains N particles and is found inside the phase volume $\mathrm{d}\boldsymbol{r}^N\mathrm{d}\boldsymbol{p}^N$ taken around the phase point $(\boldsymbol{r}^N,\boldsymbol{p}^N)$.

Exercise 4.9. Combine (2.202), (4.127), and (4.129) to deduce Gibbs's entropy formula. Do the same using (4.131). ///

In arriving at these results, we assumed that the number of mechanical degrees of freedom of system B is much larger than that of system A. This implies that the upper limit of the summation over N should be kept sufficiently small compared to the total number of particles in the composite system. However, the surroundings can be made arbitrarily large. Moreover, for a finite system volume V, $\overline{\Omega}(E,N)$ becomes negligibly small with increasing N unless E is extremely large, and the Boltzmann factor kills off the integrand. Thus, it is customary to set the upper limit of the summation to ∞. Equation (4.128) may be regarded as a *discrete* version of the Laplace transform of the canonical partition function Z.

Let us look at the other end. In principle, system A can become empty. So, the lower limit must be set to $N = 0$. What is the value of the summand in this case? If we just look at (4.130), the answer is unclear because there is no variable with respect to which to perform the integration. Let us step back a little and note that

$$\int p(E,0)\mathrm{d}E \tag{4.132}$$

is the probability that system A is empty, which, as we have just noted, can be nonzero. On the other hand, we expect that $p(E,0)$ should be zero unless E is equal to zero (or whatever any other value we assign to the vacuum):

$$p(E,0) = 0 \quad \text{if } E \neq 0\,. \tag{4.133}$$

These two equations mean that $p(E,0)$ must be proportional to the δ-function $\delta(E)$. Now, it is quite sensible to demand that, when considering system A alone, the number of microstates accessible to A when it is empty should be just one. That is, once the system boundary is fixed, there should be only one kind of vacuum:

$$\int \overline{\Omega}(E,0)\mathrm{d}E = 1\,. \tag{4.134}$$

We see from (4.127) that $\overline{\Omega}(E,0)$ is proportional to $p(E,0)$, and hence to $\delta(E)$. From (4.134), we conclude that

$$\overline{\Omega}(E,0) = \delta(E)\,, \tag{4.135}$$

which can be used in the $N = 0$ term in (4.128) to yield

$$e^{\beta\mu\cdot 0}\int \overline{\Omega}(E,0)e^{-\beta E}\mathrm{d}E = 1\,. \tag{4.136}$$

So, the $N = 0$ term is unity. Because (4.130) was obtained from (4.128), this conclusion holds for the $N = 0$ term in (4.130) as well.

Example 4.4. Ideal gas: Let us find the partition functions Ξ for a pure ideal gas, defined as a system of noninteracting identical particles. From (4.130),

$$\Xi = \sum_{N=0}^{\infty} \frac{e^{\beta\mu N}}{h^{3N}N!}\int e^{-\beta H}\mathrm{d}\boldsymbol{r}^N\mathrm{d}\boldsymbol{p}^N = \sum_{N=0}^{\infty}\frac{V^N e^{\beta\mu N}}{\Lambda^{3N}N!} = \sum_{N=0}^{\infty}\frac{1}{N!}\left(\frac{Ve^{\beta\mu}}{\Lambda^3}\right)^N\,. \tag{4.137}$$

Using (B.5), we find

$$\Xi = \exp\left[\frac{Ve^{\beta\mu}}{\Lambda^3}\right]\,. \tag{4.138}$$

The quantity

$$z := \frac{e^{\beta\mu}}{\Lambda^3} \tag{4.139}$$

occurring in (4.138) is referred to as the **absolute fugacity**. For an ideal gas, this quantity is just the number density of particles:

$$z = \frac{N}{V} \quad \text{(ideal gas only)}. \tag{4.140}$$

To see this, we recall (2.202) and the Euler relation (2.148) to obtain

$$\Omega = -PV \quad \text{(homogeneous system only)}. \tag{4.141}$$

Since (2.148) holds only for a homogeneous body, the same restriction applies to this identity. Of course, an instantaneous configuration of particles will never be homogeneous. But, in writing thermodynamic identity such as (4.141), we are concerned only with the average behavior over a long duration of time. When interaction between the particles and the system wall can be ignored, and if there is no external field, all points in the system are equivalent and a fluid phase will be homogeneous, again upon long-time averaging.

Using (4.138),

$$-PV = -k_BT\ln\Xi = -k_BTzV\ . \tag{4.142}$$

Recalling the ideal gas equation of state, $PV = Nk_BT$, we arrive at (4.140). For an explicit derivation of the equation of state, see Exercise 4.11.

Example 4.5. Probability of a cavity formation: A gas consisting of a large number (N) of noninteracting identical particles is held at a constant temperature T and occupies a macroscopic volume V. If we focus on a very small region of space of volume v taken inside V, where $v \ll V$, the number of particles in v will fluctuate with time. Let us find the probability p_0 that v contains no particle at all.

Approach 1: The Hamiltonian of the macroscopic V may be written as the sum of two Hamiltonians one pertaining only to the particles in v and the other only to $V - v$. This implies that the interaction between regions v and $V - v$ is sufficiently weak. Because V is much larger than v, region $V - v$ acts as a reservoir of the energy and the particles for region v, allowing us to describe region v by a grand canonical ensemble. The relevant partition function is

$$\Xi = \sum_{n=0}^{\infty} e^{\beta\mu n}\int \overline{\Omega}(E,n)e^{-\beta E}\mathrm{d}E = \exp\left[\frac{ve^{\beta\mu}}{\Lambda^3}\right]\ . \tag{4.143}$$

We recall that the n-th term in the summation is the unnormalized probability of finding n particles in v and that Ξ is the normalization constant.

The desired probability, therefore, is obtained by dividing the $n = 0$ term by Ξ. But, the term in question is just one. So,

$$p_0 = \frac{1}{\Xi} = \exp\left[-\frac{ve^{\beta\mu}}{\Lambda^3}\right] . \tag{4.144}$$

Equation (4.142) applied to region v gives

$$-Pv = -k_BT \ln \Xi , \tag{4.145}$$

Thus,

$$p_0 = \frac{1}{\Xi} = e^{-Pv/k_BT} , \tag{4.146}$$

in which we recognize Pv as the reversible work required to create a cavity of volume v in a gas held at the pressure P. This result is an example of (2.180).

Approach 2: We can treat the entire gas phase in V using a canonical ensemble, for which the relevant partition function is

$$Z = \frac{1}{N!}\frac{V^N}{\Lambda^{3N}} , \tag{4.147}$$

where N is the total number of particles in V. Only a subset of all microstates embraced by Z is consistent with the condition that v being empty. The partition function computed under this condition is

$$Z_0 = \frac{1}{N!}\frac{(V-v)^N}{\Lambda^{3N}} . \tag{4.148}$$

Thus,

$$p_0 = \frac{Z_0}{Z} = \left(1 - \frac{v}{V}\right)^N \tag{4.149}$$

Let us take the limit of $N \to \infty$ while holding $n_v := N/V$ constant. This is known as the **thermodynamic limit**. With the help of (B.24), we see that p_0 approaches

$$e^{-n_v v} \tag{4.150}$$

in this limit. But, this is just (4.146) because of the ideal gas equation of state. Note that the use of grand canonical ensemble in Approach 1 implies the thermodynamic limit for region $V - v$.

Exercise 4.10. For Ξ given by (4.128),

a. Show that

$$\langle N \rangle = \left(\frac{\partial \ln \Xi}{\partial \beta\mu}\right)_{\beta,V} . \tag{4.151}$$

b. Show that

$$\langle N^2\rangle - \langle N\rangle^2 = \left(\frac{\partial^2 \ln \Xi}{\partial \beta\mu^2}\right)_{\beta,V} . \tag{4.152}$$

The notation is a bit clumsy here, but the derivative is twice with respect to $\beta\mu$.

c. Show that

$$\left(\frac{\partial^2 \ln \Xi}{\partial \beta\mu^2}\right)_{\beta,V} = \frac{k_B T \kappa_T \langle N\rangle^2}{V} , \tag{4.153}$$

where

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_{T,N} \tag{4.154}$$

is the **isothermal compressibility**. ///

Combining (4.152) and (4.153), we obtain

$$\frac{\langle N^2\rangle - \langle N\rangle^2}{\langle N\rangle^2} = \frac{k_B T n_v \kappa_T}{\langle N\rangle} , \tag{4.155}$$

where $n_v := \langle N\rangle/V = 1/\underline{V}$. But, since $k_B T n_v \kappa_T$ is an intensive quantity, (4.155) implies that

$$\frac{\Delta_{\text{rms}} N}{\langle N\rangle} \sim \sqrt{\frac{1}{\langle N\rangle}} , \tag{4.156}$$

which is just another example of (4.15).

Exercise 4.11. In this chapter, we found expressions for partition functions C_M, Z, Y, and Ξ of a pure ideal gas. Assuming that N is large, find the relationships among:

a. T, U, and N
b. T, P V, and N
c. $e^{\mu/k_B T}$, V, N, and Λ

for each ensemble. Verify that, for large N, the results are independent of the choice of the ensemble. (Omit Y if you skipped Sect. 4.6.) ///

Exercise 4.12. So far, we discussed canonical, isothermal–isobaric (Sect. 4.6),and grand canonical ensembles. They are suitable for describing a system held at constant (T,V,N), (T,P,N), and (T,V,μ), respectively. In this problem, you are invited to explore the possibility of constructing a statistical ensemble suitable for a system held at constant (T,P,μ).

By a simple generalization of the expressions for Z, Y, and Ξ, it seems reasonable to expect that the partition function X for this ensemble is given by

$$X(T,P,\mu) = \int_0^\infty e^{-\beta PV} \sum_{N=0}^{\infty} \frac{e^{\beta\mu N}}{h^{3N} N!} \int e^{-\beta H(\boldsymbol{r}^N,\boldsymbol{p}^N)} \mathrm{d}\boldsymbol{r}^N \mathrm{d}\boldsymbol{p}^N \mathrm{d}V \tag{4.157}$$

in the case of a classical system of identical particles.

a. What is the probability $\rho(\boldsymbol{r}^N, \boldsymbol{p}^N, V, N)\mathrm{d}\boldsymbol{r}^N \mathrm{d}\boldsymbol{p}^N \mathrm{d}V$ of finding the system with N particles, with volume in the interval $(V, V + \mathrm{d}V]$, and within an infinitesimal volume element $\mathrm{d}\boldsymbol{r}^N \mathrm{d}\boldsymbol{p}^N$ taken around the phase point $(\boldsymbol{r}^N, \boldsymbol{p}^N)$?

b. Using Gibbs's entropy formula, evaluate the *numerical value* of $-k_B T \ln X$. Is this ensemble useful at all? Why or why not?

c. Suppose that the system is an ideal gas. Does X for the ideal gas agree with what you found in part b? If not, provide a possible explanation for the disagreement.

///

4.8 Frequently Used Symbols

$\langle A \rangle$, ensemble average of a dynamical variable A.

f , the number of mechanical degrees of freedom.
h , Planck constant. 6.626×10^{-34}(J·s).
$\hbar$, $h/2\pi$.
k_B , Boltzmann constant, 1.3806×10^{-23} J/K.
m_i , mass of the ith particle.
p_i , generalized momentum conjugate to q_i.
p^f , collective notation for $p_1, \dots, p_f$.
$\boldsymbol{p}_i$, linear momentum of the ith particle.
$\boldsymbol{p}^N$, collective notation for $\boldsymbol{p}_1, \dots, \boldsymbol{p}_N$.
q_i , the ith generalized coordinate.
q^f , collective notation for $q_1, \dots, q_f$.
$\boldsymbol{r}_i$, position vector of the ith particle.
$\boldsymbol{r}^N$, collective notation for $\boldsymbol{r}_1, \dots, \boldsymbol{r}_N$.
$\boldsymbol{v}_i$, velocity vector of the ith particle.
$\boldsymbol{v}^N$, collective notation for $\boldsymbol{v}_1, \dots, \boldsymbol{v}_N$.
z , absolute fugacity. $e^{\beta\mu}/\Lambda^3$.

A , a generic dynamical variable.
C_M , microcanonical partition function.
E , energy of the system.
F , Helmholtz free energy.
H , Hamiltonian.
N , total number of particles in a system.
P , pressure.
S , entropy.
T , absolute temperature.
U , internal energy.
V , volume.
Y , isothermal–isobaric partition function.

Z , canonical partition function.

$\mathscr{P}$, the number of permutations of identical particles.
$\mathscr{W}(E)$, the number of microstates with $H \leq E$.

β , $1/k_BT$.
$\delta(x)$, Dirac δ-function.
ρ , statistical weight.
μ , chemical potential.
$\theta(x)$, step function defined by (D.2).

Λ , thermal wavelength $h/\sqrt{2\pi m k_B T}$ of a particle of mass m.
Ξ , grand canonical partition function.
$\overline{\Omega}$, density of states.

References and Further Reading

1. Hill T L (1986) An Introduction to Statistical Thermodynamics. Dover, New York
   See Chap. 8 for a more detailed treatment of diatomic molecules. Chap. 9 is on polyatomic molecules.
2. McQuarrie D A (2000) Statistical Mechanics. University Science Books, Sausalito, California