Chapter 3

Equilibrium Statistical Mechanics

Chapter 1 and Chapter 2 have provided us with qualitative and quantitative reasons to accept equilibrium statistical mechanics as a relevant description of the late time statistics of isolated systems. In this Chapter, we study the properties of the corresponding ensembles and show their equivalence in the large system size limit. This allows us to recover standard results of thermodynamics.

3.1 The microcanonical ensemble

For an isolated system at energy E, all states with the same energy E are assumed equiprobably in the steady state. Formally, if we denote by $\{\mathcal{C}\}$ the configurations of the system, then

$$p(\mathcal{C}) = \begin{cases} \frac{1}{\Omega(E)}, & \text{if } E(\mathcal{C}) = E.\\ 0, & \text{otherwise.} \end{cases}$$
(3.1)

We define the microcanonical entropy as

$$S_m(E) = k_B \ln \Omega(E) \tag{3.2}$$

and the microcanonical temperature as

$$\frac{1}{T_m(E)} = \frac{\partial S_m(E)}{\partial E} \tag{3.3}$$

The heat capacity quantifies how much energy δE is needed to increase the temperature by δT . It is defined as $\frac{1}{c_V} \sim \frac{\delta T}{\delta E}$. Inverting T(E) into E(T), one obtains:

$$c_V = \frac{\partial E}{\partial T} \,. \tag{3.4}$$

The subscript V refers to the fact that, since we are in the microcanonical ensemble, the value V is kept constant as $E \to E + \delta E$ and $T \to T + \delta T$.

3.1.1 Continuous systems

Density of states. If we consider a system described by continuous variables, defining a "number" $\Omega(E)$ does not make sense. Instead, we aim at characterizing the "measure" of the energy surface defined by $E(\mathcal{C}) = E$. If we consider any measure that is absolutely continuous with respect to the Lebesgue measure in the full space, the measure of this surface is zero! Instead, a solution is to consider energy configurations $E \leq E(\mathcal{C}) \leq E + \delta E$. Then, the volume/measure of this set of points scales as $\Omega(E, \delta E) \simeq \omega(E)\delta E$, where $\omega(E)$ is called the density of states of the system.

The entropy is then given by $S = k_B \ln \Omega = k_B \ln \omega(E) + k_B \ln \delta E$. In the large system size limit, $k_B \ln \omega(E)$ diverges, and the constant term $k_b \ln \delta E$ can be neglected.

Phase space measure. The average of a continuous observable O(x) with respect to a probability measure μ can be written as

$$\langle O \rangle = \int O(x)\mu(\mathrm{d}x) = \int O(x)\mathrm{d}\mu(x) = \int O(x)p(x)\mathrm{d}x , \qquad (3.5)$$

where we have introduce the probability density p(x). If we make a change of variable $x \to y = \lambda x$, which amounts in changing the units in which we measure x^1 , then

$$dy = \lambda dx$$
 and $p(y) = p(x)\frac{dx}{dy} = \frac{p(x)}{\lambda}$, (3.6)

so that p(y)dy = p(x)dx and the value of the average in Eq. (3.5) does not change. When we change the units of the observable, the measure and the probability distribution change in opposite ways so as to leave the probability measure unchanged.

Now, consider a classical system, whose probability measure is

$$\underbrace{d\rho_E(\{\vec{q_i}, \vec{p_i}\})}_{\text{probability measure}} \equiv \underbrace{\frac{1}{\Omega(E, \delta E)} \mathbbm{1}_{H(\{\vec{q_i}, \vec{p_i}\}) \in [E, E+dE]}}_{\text{probability density}} \underbrace{\prod_i d^3 \vec{q_i} d^3 \vec{p_i}}_{\text{Phase space measure}} (3.7)$$

Here also, $d\rho$ is independent on the choice of units: if we change units, both the phase space volume $\Omega(E)$ and the measure $d^3 \vec{q_i} d^3 \vec{p_i}$ change in the same way, effectively compensating. Conversely, the entropy $S = k_B \ln \Omega(E)$ depends on the unit, which is unfortunate. To address this, we introduce a reference phase space volume h^{3N} , where h is the Planck constant, such that:

$$d\rho_E(\{\vec{q_i}, \vec{p_i}\}) = \frac{h^{3N}}{\Omega(E, \delta E)} \mathbb{1}_{H(\{\vec{q_i}, \vec{p_i}\}) \in [E, E+dE]} \prod_i \frac{d^3 \vec{q_i} d^3 \vec{p_i}}{h^{3N}}$$
(3.8)

This make our measurement of the volume occupied by states of energy E independent of the units, so that $S = k_B \ln \frac{\Omega(E)}{h^{3N}}$ also does not depend on the the units. Moreover, as we will see in Chapter 5, the choice of h^{3N} is consistent with the high temperature limit of quantum statistical mechanics.

3.1.2 The ideal gas

Consider N particles in a box of volume $V = L^3$. In the dilute limit, their interactions suffice for the gas to equilibrate but they are typically small compared to the kinetic energy of the particles. We thus approximate the energy as $H = \sum_{i=1}^{N} \frac{\vec{p}^2}{2m}$.

Let us first compute the normalized density of states $\tilde{\Omega}(E, \delta E) = \Omega(E, \delta E)/h^{3N}$:

$$\tilde{\Omega}(E,\delta E) = \int_{E \le H(\{\vec{q}_i, \vec{p}_i\}) \le E + \delta E} \prod_{i=1}^N \frac{d^3 \vec{q}_i d^3 \vec{p}_i}{h^3}$$
(3.9)

$$= \left(\frac{V}{h^3}\right)^N \int_{E \le \sum_i \frac{\vec{p}_i^2}{2m} \le E + \delta E} \prod_{i=1}^N d^3 \vec{p}_i$$
(3.10)

$$= \left(\frac{V}{h^3}\right)^N \left[v(E+\delta E) - v(E)\right]$$
(3.11)

where v(E) is the volume of the hypersphere such that $\sum_{i} \frac{\vec{p_i}^2}{2m} \leq E$. It follows that $\tilde{\Omega}(E, \delta E) \simeq \frac{\delta E}{h^{3N}} \frac{dv(E)}{dE}$. Let's compute v(E):

$$v(E) = L^{3N} \int_{\sum_{i} \vec{p}_{i}^{2} \le 2mE} \prod_{i=1}^{N} d^{3} \vec{p}_{i} = L^{3N} \int_{\sum_{i} x_{i}^{2} \le 2mE} \prod_{i=1}^{3N} d^{3} x_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \le 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \ge 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \ge 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \ge 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \ge 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \ge 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \ge 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \ge 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \ge 1} \prod_{i=1}^{3N} du_{i} = L^{3N} (2mE)^{3N/2} \int_{\sum_{i} u^{2} \ge 1} \prod_{i=1}^{3N}$$

¹For instance, if x is a length in meters and y the same length in millimeters, then $\lambda = 10^3$.

where in the last step we have changed variables $x_i = \sqrt{2mE}u_i$. In this way, we have make the integral dimension-free: all the dimensions now enter in the prefactor. By going to spherical coordinates, we get:

$$v(E) = L^{3N} (2mE)^{3N/2} \Omega_{\rm so}(3N) \int_0^1 \mathrm{d}u e^{3N-1} = L^{3N} \frac{(2mE)^{3N/2}}{3N} \Omega_{\rm so}(3N)$$
(3.12)

where $\Omega_{\rm so}(3N)$ is the total solid angle in 3N-dimensions.

Solid angle in d-dimensions. To determine $\Omega_{so}(3N)$, we can try to find an integral that we can do both in cartesian and spherical coordinates. A standard possibility is to use Gaussian integrals. We thus introduce

$$I_{d} \equiv \left(\underbrace{\int_{-\infty}^{+\infty} dx e^{-x^{2}}}_{\sqrt{\pi}}\right)^{d} = \pi^{d/2} = \int \prod_{i=1}^{d} dx_{i} e^{-\sum_{i} x_{i}^{2}} = \Omega_{\rm so}(d) \int_{0}^{\infty} dr r^{d-1} u^{-r^{2}}$$
(3.13)

where in the last integral, we have gone to spherical coordinates and introduced $r^2 = \sum_i x_i^2$. Using the change of variable $\omega = r^2$, we find

$$I_d = \frac{1}{2}\Omega_{\rm so}(d)\int_0^\infty d\omega e^{-\omega}\omega^{d/2-1} = \frac{1}{2}\Omega_{\rm so}(d)\Gamma\left(\frac{d}{2}\right)$$
(3.14)

where $\Gamma(u) = \int_0^\infty d\omega \omega^{u-1} e^{-\omega}$ is the Gamma function such that $\Gamma(n) = (n-1)!$ for n a positive integer. Comparing Eq. (3.13) and Eq. (3.14) one obtains for the solid angle in d-dimensions,

$$\Omega_{\rm so}(d) = \frac{2\pi^{d/2}}{(\frac{d}{2} - 1)!} , \qquad (3.15)$$

and thus

$$v(E) = L^{3N} (2mE)^{3N/2} \frac{2\pi^{3N/2}}{3N(\frac{3N}{2} - 1)!}, \qquad (3.16)$$

where we extend the factorial notation to half integers, to lighten our notations. Recalling that $\tilde{\Omega}(E) = v'(E) \frac{\delta E}{\hbar^{3N}}$, one obtains for the density of states:

$$\tilde{\Omega}(E) = \delta E \left(\frac{L}{h}\right)^{3N} m(2mE)^{3N/2-1} \frac{2\pi^{3N/2}}{\left(\frac{3N}{2} - 1\right)!}$$
(3.17)

We can now use Eq. (3.17) to compute the microcanonical entropy:

$$S_m = k_B \ln \tilde{\Omega}(E) \simeq 3N k_B \ln \left(\frac{L}{h}\right) + \frac{3N}{2} k_B \ln \left(2mE\pi\right) - k_B \ln \left[\left(\frac{3N}{2} - 1\right)!\right] + \mathcal{O}(N) , \qquad (3.18)$$

where the last term, $\mathcal{O}(N)$ includes a bunch of subdominant term that involve, for instance, δE . Let us try to simplify further the expression of S_m , by using

Stirling formula: $n! \sim \sqrt{2\pi n \left(\frac{n}{e}\right)^n} \Rightarrow \ln n! \simeq n \ln \left(\frac{n}{e}\right) + o(n)$ where a = o(n) if $\frac{a}{n} \xrightarrow[n \to \infty]{} 0$. Using Stirling formula to expand the factorial in Eq. (3.18) we obtain

$$S_m \simeq Nk_B \ln\left[V\left(\frac{4mEe\pi}{3Nh^2}\right)^{3/2}\right]$$
(3.19)

A number of comments are in order:

	m	ð	Ś	\wedge	m	Á
	0	V	3	C)	3	9
		<u>(</u>		<u>@</u>		Ø
	(1)		<₹		()	
-	V					

Figure 3.1: If we can label our particles, they are distinguishable and all these configurations are different. They should thus all be included in the computation of $\Omega(E)$. If the particles are undistinguishable, all these configurations are identical and we have overestimated $\Omega(E)$ by a factor of N!.

- The fact that $S_m \to \infty$ as $N \to \infty$ legitimates disregarding the term δE .
- S_m increases with E as expected. This grants the positivity of the temperature of the ideal gas $\frac{1}{T} = \frac{\partial S}{\partial E} > 0.$
- S_m is not extensive. One can easily check that $S_m(\lambda E, \lambda V, \lambda N) \neq \lambda S_m(E, V, N)$. While we have a factor E/N in S_m , that is intensive, the isolated factor of V is not.

Indistinguishibility: The super-extensivity of the ideal-gas entropy calculated above indicates that there is an issue with our computations. Indeed, for an ideal gas of indistinguishable particles, we expect the entropy to be extensive from thermodynamics consideration. This was first discussed by Gibbs and is (unfortunately) called the Gibbs paradox in the context of the mixing of two gases.

Let us first note that the computation is perfectly correct if particles are distinguishable, i.e. if we can label them and track them, as illustrated in Fig. 3.1. This is what our computation did: we gave a position q_i to particle *i* and counted all $\{q_i\}$ as distinct configurations. For indistinguishable particles, however, we have overcounted configurations since we have counted as different states configurations of particles with identical positions and permutations of labels. This is wrong since $\{q_i\}$ and $\{q_{\sigma(i)}\}$ —with σ a permutation of $\{1, \ldots, N\}$ —do not correspond to different configurations when the particles are indistinguishable.

To sum up, if particles are distinguishable, permutations of particle labels create new configurations and the entropy is indeed superextensive. On the other hand, if the particles are indistinguishable, swapping particle labels do not create new configurations and we have to correct our estimate of $\Omega(E)$ by dividing it by a factor of N!. Then the entropy becomes

$$S_m(E) \simeq Nk_B \ln\left[\frac{eV}{N} \left(\frac{4\pi emE}{3Nh^2}\right)^{3/2}\right].$$
(3.20)

The entropy of the ideal gas is now, as expected, extensive.

Intensive thermodynamics variables. In the infinite system limit, in which thermodynamic hold, S_m is ... infinite! It is thus often more practical to work with its intensive counterpart. From Eq. (3.20), one defines the entropy density as $s = \frac{S_m}{N}$, which is a function of the free volume per particle, $v = \frac{V}{N}$, and the energy density, $e = \frac{E}{N}$, that are both intensive quantities.

Comment: In quantum mechanics, particles are waves and cannot be distinguished if they are of the same nature. They should thus be treated as indistinguishable.



Figure 3.2: Sketch of the entropy of a two-level systems, highlighting the positive and negative temperature regions.

Thermodynamics quantities: By taking the log of $\Omega(E)$, i.e. the normalization factor of the microcanonical probability distribution, we have determined the entropy, which is the thermodynamic potential of the microcanonical ensemble. From its knowledge, we can now determine other thermodynamic observables. This pattern will happen again in all the ensembles that we will construct.

We thus use Eq. (3.20) to compute the thermodynamic quantities for the ideal gas. From $\frac{1}{T_m} = \frac{\partial S_m}{\partial E} = \frac{3}{2E}Nk_B$, we get

$$E = \frac{3}{2}Nk_BT_m \tag{3.21}$$

$$c_V = \frac{\partial E}{\partial T_m} = \frac{3}{2}Nk_B \tag{3.22}$$

3.1.3 Discrete systems: the two-level system

In many systems, the variations of energy are not (solely or at all) due to motion in space, but instead are due to changes in discrete observables. An important example is that of localized electrons on a lattice i the presence of a magnetic field. Taking into account the g ratio of the electrons, their energy is $E = -\mu h \sum_{i=1}^{N} \sigma_i - J \sum_{iVj} \sigma_i \sigma_j$ where μ is the magnetic moment of the electrons, J is the exchange energy, $\sigma_j = \pm 1$ is a normalized spin, and iVj refers to a sum over nearest neighbors, which is also often noted $\langle i, j \rangle$. We will come back to this important example in Chapter 4, when we study the Ising model. For now, we discuss the simplest of discrete systems.

Two level systems: We consider N non-interacting atoms on a lattice that can be in two energy levels, with energy 0 or ε . Then $H = \sum_{i=1}^{N} \varepsilon_i = n\varepsilon = pN\varepsilon$, where $n \in \{0, \dots, N\}$ and p is the fraction of atoms with energy ε . Since the atoms have fixed positions, they are distinguishable and

$$\Omega(E = pN\varepsilon) = \binom{N}{Np} = \frac{N!}{(pN)!(N - pN)!}$$
(3.23)

which is the number of ways to choose pN atoms among N. We use Eq. (3.23) together with Stirling formula to compute the entropy,

$$S_m = k_B \ln \Omega(E) \simeq -k_B N \left[p \ln p + (1-p) \ln(1-p) \right]$$
(3.24)

with $p = \frac{E}{N\varepsilon}$. $S_m(E)$ is sketched in Fig. 3.2.



Figure 3.3: Sketch of the relationship between the temperature and the energy of a two-level systems.

We see that $S_m \to 0$ both for $p \to 0$ and $p \to 1$, since in both cases there is only one possible configuration corresponding respectively to E = 0 or E = N. As a consequence, S_m is not a monotonous function, and the temperature can be negative. In particular,

$$\frac{1}{T_m} = \frac{\partial S}{\partial E} = \frac{\partial S_m}{\partial p} \frac{\partial p}{\partial E} = \frac{k_B}{\varepsilon} \ln\left(\frac{1-p}{p}\right) < 0 \quad \text{if} \quad p > \frac{1}{2}$$
(3.25)

The positive-T region corresponds to $E < N\varepsilon/2$ while the negative-T region corresponds to $E > N\varepsilon/2$, as shown in Fig. 3.3.

Inverting this relation and using $e^{\varepsilon/k_BT} = \frac{1-p}{p}$, one obtains for the energy:

$$E = \frac{N\varepsilon}{1 + e^{\varepsilon/k_B T}} \tag{3.26}$$

and for the heat capacity:

$$c_V = \frac{\mathrm{d}E}{\mathrm{d}T} = Nk_B \frac{\varepsilon^2}{(k_B T)^2} \frac{e^{\varepsilon/k_B T}}{(1 + e^{\varepsilon/k_B T})^2}$$
(3.27)

We observe an exponential decay of the heat capacity in the limit $T \to 0$, which is typical of systems with an energy gap (as opposed to the continuous values of H for a classical ideal gas).

3.1.4 Large systems and thermodynamics

3.1.4.1 Macrostates

As for the kinetic theory of gases, the microscopic configurations of the system, referred to as microstates C_m , often contain too much information on the system. One may then be more interested into coarsegrained descriptions of the system. A standard way to proceed is to group microstates together into macrostates C_M , based on some common property of interest.

For example, in a spin systems, a microstate C_m is defined by the values (S_1, \ldots, S_N) of spins. A macrostate $C_M(m)$ can be defined by the system magnetization

$$m = \frac{1}{N} \sum_{i=1}^{N} S_i$$
 (setting $\mu = 1$). (3.28)

 $\mathbf{43}$



Figure 3.4: Schematic representation of a system S in the microcanonical ensemble, divided between two subsystems S_1 and S_2 that can exchange energy. The volume and particle number of both subsystems are fixed and, for simplicity, we choose them equal: $V_1 = V_2 = V$, $N_1 = N_2 = N$.

The macrostate is then defined as

$$\mathcal{C}_M(m) = \{\mathcal{C}_m, \text{ such that } \frac{1}{N} \sum_i S_i = m\}$$
(3.29)

There are 2^N microstates C_m and only N+1 macrostates C_M , since $Nm \in \{-N, -N+2, -N+4, \ldots, N-2, N\}$. It represents a strong dimensional reduction, which justifies the terminology "micro" vs "macro". By construction, the probability of a macrostate is

$$P(\mathcal{C}_M) = \sum_{\mathcal{C}_m \in \mathcal{C}_M} P(\mathcal{C}_m) .$$
(3.30)

Since there are typically many C_m 's within one macrostate C_M , we can often simplify $P(C_M)$ to extract meaningful information on the macrostates. This will be an example of the concept that we discussed in the introduction: More is simpler.

An important example where the use of macrostate plays an important role is the phenomenon of *equilibration*, when two systems are put into contact.

3.1.4.2 Sub-systems and equilibration

Consider an isolated system S divided into two subsystems S_1 and S_2 that can exchange energy, as in Fig. 3.4, so that the any microstate of the full system can be written as $C = C_1 \otimes C_2$. We assume the subsystems to be sufficiently large such that the interaction energy can be neglected and the total energy can be expressed simply as the sum of the energy of the two subsystems, $E \simeq E_1 + E_2$. We denote $\Omega_1(E_1)$ the number of configurations C_1 of S_1 with energy E_1 and $\Omega_2(E_2)$ the number of configurations C_2 of S_2 with energy E_2 . Because the full system is isolated, the probability to observe a configuration C of the total system is given by the microcanonical measure

$$P_E(\mathcal{C}) = P(\mathcal{C}_1, \mathcal{C}_2) = \frac{1}{\Omega(E(\mathcal{C}))} \delta(E(\mathcal{C}) - E)$$
(3.31)

where $\Omega(E) = \sum_{E_1} \Omega_1(E_1) \times \Omega_2(E_2 = E - E_1).$



Figure 3.5: Sketch of the variation of the entropy density $s_1(E_1) + s_2(E - E_1)$. The integral in Eq. (3.37) is dominated by the vicinity of E_1^* , which maximizes the entropy density of the macrostate of energy E_1 .

Typical values of E_1 and E_2 . We want to use Eq. (3.31) to characterize the fluctuations of E_1 and $E_2 = E - E_1$. To do so, we consider the corresponding macrostate given by all the microstates (C_1, C_2) such that $E(C_1) = E_1$. The probability of this macrostate is given by the sum of the probabilities of the microstates it contains:

=

$$\underbrace{P(E_1, E_2 = E - E_1)}_{P(E_1)} = \sum_{\mathcal{C}_1, \mathcal{C}_2 | E_1 + E_2 = E} \frac{1}{\Omega(E_1 + E_2)}$$
(3.32)

$$= \frac{1}{\Omega(E_1 + E_2)} \sum_{\mathcal{C}_1 | E(\mathcal{C}_1) = E_1} \sum_{\mathcal{C}_2 | E_2 = E - E_1} 1$$
(3.33)

$$= \frac{1}{\Omega(E_1 + E_2)} \sum_{\mathcal{C}_1 | E(\mathcal{C}_1) = E_1} \Omega_2(E - E_1)$$
(3.34)

$$=\frac{\Omega_1(E_1) \times \Omega_2(E_2)}{\Omega(E)} \tag{3.35}$$

Let us now show that $P(E_1)$ is peaked at a typical value E_1^* . In large systems, we replace \sum_{E_1} with an integral, assuming E varies smoothly, and write

$$\Omega(E) \simeq \int dE_1 \Omega_1(E_1) \Omega_2(E - E_1)$$
(3.36)

Using the definition of microcanonical entropy, this becomes

$$\Omega(E) = \int dE_1 e^{\frac{S_1(E_1)}{k_B} + \frac{S_2(E-E_1)}{k_B}} = \int dE_1 e^{\frac{N}{k_B}[s_1(E_1) + s_2(E-E_1)]}$$
(3.37)

where in the last step we introduced the intensive entropy densities $s_i \equiv \frac{S_i}{N} \underset{N \to \infty}{\sim} \mathcal{O}(1)$. In the large N limit, the integral is strongly dominated by the maximum E_1^* of $\phi(E_1) = s_1(E_1) + s_2(E - E_1)$, as sketched in Fig. 3.5. Let us now show how to compute the corresponding value of $\Omega(E)$.

Saddle point approximation

We want to approximate the integral $I = \int_a^b dx e^{N\phi(x)}$ as $N \to \infty$ when $\phi(x)$ is a real function with a



Figure 3.6: In the large N limit, only the vicinity of x_0 plays a role in the integral of $\varphi_N(x) = e^{N\phi(x)}$.

maximum at $x = x_0 \in (a, b)$ and $\phi(x) \underset{N \to \infty}{\sim} \mathcal{O}(1)$. Taking a Taylor expansion around the maximum:

$$I = \int_{a}^{b} dx e^{N\left[\phi(x_{0}) + (x - x_{0}) \underbrace{\phi'(x_{0})}_{=0} + \frac{1}{2}(x - x_{0})^{2} \underbrace{\phi''(x_{0})}_{<0} + \dots\right]}$$
$$\simeq e^{N\phi(x_{0})} \int_{a}^{b} dx e^{-\frac{N}{2}(x - x_{0})^{2}|\phi''(x_{0})|} = e^{N\phi(x_{0})} \sqrt{\frac{2\pi}{N|\phi''(x_{0})|}}$$

In the last step, we have used that only the vicinity of x_0 plays a role in the integral of $\varphi_N(x) = e^{N\phi(x)}$ so that we can replace the limits [a, b] by $(-\infty, \infty)$, as sketched in Fig.

Comment: We could expand the terms neglected in the first line to obtain the series of moments of a Gaussian distribution. One then check that, in the limit $N \to \infty$, these terms are subdominant.

By applying the saddle point approximation to $\Omega(E)$, one then finds²:

$$\Omega(E) \underset{N \to \infty}{\approx} e^{\frac{N}{k_B} \left[s_1(E_1^* + s_2(E - E_1^*)) \right]}$$
(3.38)

where E_1^* is such that $\phi'(E_1) = 0$. Multiplying this condition by N leads to the following saddle-point condition for E_1^* :

$$S_1'(E_1^*) - S_2'(E - E_1^*) = 0. (3.39)$$

Using Eq. (3.38) into Eq. (3.32) we then find that the probability distribution of E_1 is given by:

$$P(E_1) = \frac{1}{\Omega(E)} e^{\frac{N}{k_B} [s_1(E_1) - s_2(E - E_1)]} \approx e^{\frac{N}{k_B} [s_1(E_1) - s_1(E_1^*) + s_2(E - E_1) - s_2(E - E_1^*)]}$$
(3.40)

One sees that $P(E_1) \xrightarrow[N \to \infty]{} 0$ if $E_1 \neq E_1^*$. In the large-size limit, the distribution of E_1 is thus sharply peaked around the typical value E_1^* .

We note that condition (3.39) states that the value of the energy E_1 that dominates the probability $P(E_1)$ is such that the subsystems S_1 and S_2 have the same microcanonical temperatures:

$$\frac{\mathrm{d}S_1}{\mathrm{d}E_1}\Big|_{E_1^*} = \frac{\mathrm{d}S_2}{\mathrm{d}E_2}\Big|_{E_2^* = E - E_1^*} \Leftrightarrow T_{m,1} = T_{m,2} .$$
(3.41)

This is sometimes referred to as the 0^{th} law of thermodynamics: Large isolated systems that are put in contact relax to a state where their temperatures are equal.

²here $a \approx b$ refers to a logarithmic equivalence, i.e. that $\ln a \underset{N \to \infty}{\sim} \ln b$ so that polynomial prefactors of a and b in N are neglected.

Since E_1^* maximises the total entropy of the isolated system, $S(E) = S_1(E_1^*) + S_2(E_2^*)$, we find that the heat capacities c_V^1 and c_V^2 are positive:

$$\frac{\partial^2 S_1}{\partial E_1^2} + \frac{\partial^2 S_2}{\partial E_2^2} < 0 \iff \frac{1}{c_V^1} + \frac{1}{c_V^2} > 0.$$
(3.42)

As we discuss below, the positivity of the heat capacity is a requirement for the thermodynamic stability of the system.

(Nonequilibrium) entropy of a macrostate. We note that $\Omega_1(E_1) \times \Omega_2(E - E_1)$ is the total number of microstates in the macrostate of energy E_1 . In the literature, it is not uncommon to see $\phi(E_1) = \frac{1}{Nk_B} \log[\Omega_1(E_1) \times \Omega_2(E - E_1)]$ called the "entropy" density of the macrostate of energy E_1 , or the "nonequilibrium entropy" density of the macrostate of energy E_1 . The derivation above then shows that, in the large-size limit, the system "maximizes the entropy of the macrostate". The "equilibrium entropy" $S(E_1^*)$ is then the largest of the "nonequilibrium entropoes" $S(E_1)$. This notation can be a bit confusing, since $S = k_B \ln \Omega(E)$ is defined from the density of state of the full system, while the "nonequilibrium entropy" is defined from the density of state of a macrostate.

Fluctuations of E_1 : Taking the Taylor expansion of the exponent in Eq. (3.40), close to $E_1 \simeq E_1^*$, then leads to a Gaussian approximation for $P(E_1)$:

$$P(E_1) \approx \exp\left[-\frac{(E_1 - E_1^*)^2}{2k_B T^2 \frac{c_V^1 c_V^2}{c_V^1 + c_V^2}}\right]$$
(3.43)

We recall that, since $C_V = \frac{\partial E}{\partial T}$, the heat capacity is extensive: $C_V \sim N$. The variance of the energy thus scales as $\langle (E_1 - E_1^*) \rangle \sim C_V \sim N$ and typical fluctuations around the mean value E_1^* scale as $\sim \sqrt{N}$. Since $E_1^* \propto N$, the relative fluctuations of E_1^* are negligible in the large-size limit.

Third law of thermodynamics: From the definition of heat capacity, one has that

$$S(T) - S(0) = \int_0^T dT' \frac{c_V}{T'} \,. \tag{3.44}$$

At low temperatures, quantum fluctuations become important and classical statistical mechanics is not valid anymore. At T = 0, the system is in the ground state so that $S = k_B \ln g$, where g is the degeneracy of the ground state. The third law of thermodynamics is that $\lim_{T\to 0} S(T) = S_0$ is a constant finite number. If g = 1, S(0) = 0. Typically, quantum fluctuations make g subextensive so that $S(0) \ll N$ and

$$S(T) \simeq \int_0^N \mathrm{d}T \frac{c_V}{T'} \,. \tag{3.45}$$

Second law of thermodynamics: Take two systems with initial energues E_1 and E_2 . Let them evolve in contact with each other until they reach a global microcanonical equilibrium. The discussion above show that they relax to E_1^* and E_2^* such that

$$S_{\text{final}} = S_1(E_1^*) + S_2(E_2^*) \ge S_{\text{initial}} = S_1(E_1) + S_2(E_2) .$$
(3.46)

This tells us that the entropy of the isolated system $S = S_1 \cup S_2$ has increased, in agreement with the second law of thermodynamics which holds in the limit $N \to \infty$.



Figure 3.7: Sketch of an isolated system in equilibrium at volume V, with a piston exerting a force F on the system.

3.1.4.3 Pressure and the 1st law of thermodynamics

Consider an isolated system maintained at a volume V using a piston, as in Figure 3.7. To keep the volume fixed, the piston exert a force $F = f_a A$ on the system, where A is the area of the right wall and f_a is the force per unit area.

Consider compressing the system so that the right wall moves by a distance dx. The corresponding work exerted on the system is $dW = f_a A dx = f_a dV$, with the system energy evolving as $E \to E + f_a dV$. The entropy then evolves as $S_m(N, V, E) \to S_m(N, V + dV, E + dE)$ with

$$dS_m = S_m(N, V + dV, E + dE) - S_m(N, V, E)$$
(3.47)

$$= \frac{\partial S_m}{\partial V} dV + dE \underbrace{\frac{\partial S_m}{\partial E}}_{1/T_m} = \frac{\partial S_m}{\partial V} dV + f_a dV \frac{1}{T}$$
(3.48)

Let us now consider that the piston exerces a force that exactly balance the force exerted by the gas on the piston. The latter does not move and the system does not evolve, so that $dS_m = 0$ and $\frac{\partial S_m}{\partial V} = -\frac{f_a}{T}$. This last equality can be seen as a force balance: at steady state, the system exerts back a force per unit area that balances f_a , which is the pressure

$$P_m = T \frac{\partial S_m}{\partial V} \,. \tag{3.49}$$

Eq. (3.48) can thus be written as

$$\mathrm{d}S_m = \frac{P_m}{T}\mathrm{d}V + \frac{\mathrm{d}E}{T_m} \tag{3.50}$$

or equivalently

$$\mathrm{d}E = T_m \mathrm{d}S_m - P_m \mathrm{d}V , \qquad (3.51)$$

which is the first law of thermodynamics.

Summary:

- We have shown how, for any finite system, we can compute $\Omega(E)$ (or the dimensionless $\tilde{\Omega}(E)$) to characterize the probability of microstates in the microcanonical ensemble.
- We can then compute S(E) in the large N limit to construct thermodynamic observables such as the heat capacity and the pressure.



Figure 3.8: Schematic representation of a system S in contact with a thermostat. The system exchanges energy ΔE with the thermostat but their respectives particle numbers and volumes are constant.

• When interested in a given observable M, we can define the corresponding macrostates $C_M = \{C_m\}$, which are collections of microstates with the same values of M. Their probability distributions are given by $P(\mathcal{C}_M) = \frac{\Omega(\mathcal{C}_M)}{\Omega(E)}$, where $\Omega(\mathcal{C}_M)$ is the number of microstates within \mathcal{C}_M . In the large system limit, $P(\mathcal{C}_M)$ can then be expressed in a simpler form thanks to a saddle point approximation in the computation of $\Omega(E)$. When the "non-equilibrium" entropy of the macrostate, $S = k_B \log \Omega(\mathcal{C}_M)$ is extensive, this leads to a form of typicality where the measure $P(\mathcal{C}_M)$ becomes concentrated around a typical value of the observable M.

Let us now show how the concepts of macrostate introduced above can be used to change ensemble.

3.2 Canonical ensemble

3.2.1 Changing ensemble: the Boltzmann weight

Canonical distribution.

In the scenario sketched in Fig. (3.8), the system S and the thermostat are isolated but exchange energy with each others. This implies that the total energy $E_{\text{tot}} = E_S + R_{\text{th}}$ is fixed, while E_S and E_{th} can fluctuate. We assume that the system and the thermostat are both sufficiently large that the interaction energy between them can be neglected. Furthermore, the thermostat is considered much larger than the system, and $E_{\text{th}} \gg E_s$. Since the total system comprising the system S and the thermostat is isolated, the probability of a microstate $\mathcal{C} = (\mathcal{C}_S, \mathcal{C}_{\text{th}})$ is given by the microcanonical distribution

$$P_{\text{tot}}(\mathcal{C}) = \frac{1}{\Omega_{\text{tot}}(E_{\text{tot}})} \delta_{E_{\text{th}}, E_S - E_{\text{tot}}}$$
(3.52)

Since we are interested in the property of the system S only, we define a macrostate of the full system as

$$\hat{\mathcal{C}}_S = \{\mathcal{C} = \{\mathcal{C}_S, \mathcal{C}_{\text{th}}\} \text{ with } \mathcal{C}_S \text{ fixed, and } \mathcal{C}_{\text{th}} \text{ such that } E_{\text{th}}(\mathcal{C}_{\text{th}}) = E_{\text{tot}} - E_S(\mathcal{C}_S)\}$$
(3.53)

Eq. (3.53) can be seen as a marginalization of the thermostat. For simplicity, we identify $P(\hat{\mathcal{C}}_S) = P(\mathcal{C}_S)$. Then,

$$P(\mathcal{C}_S) = \sum_{\mathcal{C} \in \hat{\mathcal{C}}_S} P_{\text{tot}}(\mathcal{C}) = \frac{1}{\Omega_{\text{tot}}(E_{\text{tot}})} \sum_{\mathcal{C}_{\text{th}} \mid E(\mathcal{C}_{\text{th}}) = E_{\text{tot}} - E(\mathcal{C}_S)} 1 = \frac{\Omega_{\text{th}}[E_{\text{tot}} - E(\mathcal{C}_S)]}{\Omega_{\text{tot}}(E_{\text{tot}})}$$
(3.54)

where $\Omega_{\rm th}(E_{\rm tot})$ in the numerator counts the number of microstates of the thermostat with energy $E_{\rm th} = E_{\rm tot} - E(\mathcal{C}_S)$. Using the definition of the microcanonical entropy, we rewrite Eq. (3.54) as

$$P_{\text{tot}}(\hat{\mathcal{C}}_S) = \frac{e^{\frac{1}{k_B}S_m^{\text{th}}(E_{\text{tot}} - E(\mathcal{C}_S))}}{\Omega_{\text{tot}}(E_{\text{tot}})} \,. \tag{3.55}$$

Since $E_{\text{tot}} \gg E(\mathcal{C}_S)$, we expand

$$S_m^{\rm th}(E_{\rm tot} - E(\mathcal{C}_S)) \simeq S_m^{\rm th}(E_{\rm tot}) - E(\mathcal{C}_S) \frac{\partial S_m^{\rm th}(E_{\rm tot})}{\partial E} = S_m^{\rm th}(E_{\rm tot}) - \frac{E(\mathcal{C}_S)}{T_m^{\rm th}} .$$
(3.56)

All in all, we find

$$P(\mathcal{C}_S) = \frac{\Omega_{\rm th}(E_{\rm tot})}{\Omega_{\rm tot}(E_{\rm tot})} e^{-\frac{E(\mathcal{C}_S)}{k_B T_m^{\rm th}}} = \frac{1}{Z} e^{-\beta E(\mathcal{C}_S)} .$$
(3.57)

Here, $\beta = \frac{1}{k_B T_m^{\text{th}}}$ is the inverse temperature (in J^{-1}) of the thermostat, and $Z = \frac{\Omega_{\text{tot}}(E_{\text{tot}})}{\Omega_{\text{th}}(E_{\text{tot}})}$ is a normalization constant, which does not depend on the system configuration and ensures normalization through $\sum_{\mathcal{C}_S} P(\mathcal{C}_S) = 1$. As such, it can be directly expressed as a function of β only, as

$$Z = \sum_{\mathcal{C}_S} e^{-\beta E(\mathcal{C}_S)} \tag{3.58}$$

Comments:

- Eq. (3.57) is called the canonical distribution or the Boltzmann weight.
- Z is called the partition function. It only depends on the thermostat only through its (inverse) temperature β .
- While systems in the microcanonical ensemble are characterized by N, E and V, the system is here controlled by N, T, V, where T is the microcanonical temperature of the thermostat.

We define the Helmholtz free energy as

$$F(N, V, T) = -k_B T \ln Z(N, V; T).$$
(3.59)

Continuous systems.

For a continuous classical systems described by N pairs of positions $\vec{q_i}$ and momenta $\vec{p_i}$, the Boltzmann weight is given by:

$$P(\{\vec{q}_i, \vec{p}_i\}) = \frac{1}{Z} e^{-\beta H(\{\vec{q}_i, \vec{p}_i\})}$$
(3.60)

$$Z = \int \prod_{i} d^{3} \vec{q_{i}} d^{3} \vec{p_{i}} e^{-\beta H(\{\vec{q_{i}}, \vec{p_{i}}\})}$$
(3.61)

$$\langle O \rangle = \frac{1}{Z} \int \prod_{i} d^{3} \vec{q_{i}} d^{3} \vec{p_{i}} O(\{\vec{q_{i}}, \vec{p_{i}}\}) e^{-\beta H(\{\vec{q_{i}}, \vec{p_{i}}\})}$$
(3.62)

To make the partition function independent of the choice of units, and to makes sure that classical statistical mechanics agree with quantum mechanics, the phase space measure and the partition function are both rescaled by h^{3N} , leading to the modified expressions:

$$P(\{\vec{q}_i, \vec{p}_i\}) = \frac{1}{\tilde{Z}} e^{-\beta H(\{\vec{q}_i, \vec{p}_i\})}$$
(3.63)

so that

$$\langle O \rangle = \frac{1}{\tilde{Z}} \int \prod_{i} \frac{\mathrm{d}^{3} \vec{q_{i}} \mathrm{d}^{3} \vec{p_{i}}}{h^{3}} O(\{\vec{q_{i}}, \vec{p_{i}}\}) e^{-\beta H(\{\vec{q_{i}}, \vec{p_{i}}\})}$$
(3.64)

with

$$\tilde{Z} = \int \prod_{i} \frac{\mathrm{d}^3 \vec{q_i} \mathrm{d}^3 \vec{p_i}}{h^3} e^{-\beta H(\{\vec{q_i}, \vec{p_i}\})} \,. \tag{3.65}$$

From a classical perspective, both formalisms are entirely equivalent. Therefore, for simplicity, we drop the ~ notation in the following.

Indistinguibility: If the N degrees of freedom describe indistinguishable particles, then, as in the microcanonical ensemble, we have overcounted the number of configurations in Z. So the final correct expression for the partition function is:

$$\tilde{Z} = \frac{1}{h^{3N}N!} \int \prod_{i} d^{3}\vec{q_{i}} d^{3}\vec{p_{i}} e^{-\beta H(\{\vec{q_{i}},\vec{p_{i}}\})}$$
(3.66)

Note, however, that the same overcounting is present in the integrals entering the average $\langle O \rangle$ in Eq. (3.67) so that the measure also has to be corrected:

$$\langle O \rangle = \frac{1}{\tilde{Z}} \int \prod_{i} \frac{1}{N!} \frac{\mathrm{d}^{3} \vec{q_{i}} \mathrm{d}^{3} \vec{p_{i}}}{h^{3}} O(\{\vec{q_{i}}, \vec{p_{i}}\}) e^{-\beta H(\{\vec{q_{i}}, \vec{p_{i}}\})} .$$
(3.67)

While the factors of h and N! enter the calculation of the thermodynamic functional (in the correct way!), they do not impact any average that we do in the canonical ensemble as long as N and the number of degrees of freedom are constant. When considering subsystems in which the number of particles can fluctuates or transition between states with different numbers of degrees of freedom are allowed, then the factors of N! and h^3 can appear at the classical level. See, e.g., exercise 4 of Pset 5, "*Equilibrium ratio* between atomic and molecular forms".

Equipartition theorem: Consider an Hamiltonian $H(x_1, \dots, x_{2N})$, where $x_1, \dots, x_N = q_1, \dots, q_N$ and $x_{N+1}, \dots, x_{2N} = p_1, \dots, p_N$. Notice that in order for Z to be finite, the Hamiltonian H must diverge for $x_i \to \infty$, $\forall i$. (This is automatically true for a gas in a box). One computes:

$$\langle x_i \frac{\partial H}{x_\ell} \rangle = \frac{1}{ZN!} \int \prod_n \frac{\mathrm{d}\vec{q}_n \mathrm{d}\vec{p}_n}{h^3} \frac{\partial H}{\partial x_\ell} e^{-\beta H} \cdot x_i \stackrel{\text{d}}{=} \frac{k_B T}{ZN! h^{3N}} \int \prod_n \frac{\mathrm{d}\vec{q}_n \mathrm{d}\vec{p}_n}{h^3} e^{-\beta H} \frac{\partial x_i}{\partial x_\ell} = k_B T \delta_{i,\ell} \qquad (3.68)$$

where we made use of the integration by part in the one-before-last equality. If $x_i = p_i$ and $x_\ell = p_\ell$ one gets:

$$\left\langle \frac{p_i p_\ell}{2m} \right\rangle = \frac{k_B T}{2} \delta_{i,\ell} \tag{3.69}$$

One recovers that each quadratic degree of freedom contributes $\frac{k_BT}{2}$ to the average energy.

3.2.2 Two examples: ideal gas & two-level systems

Ideal gas: Consider N noninteracting indistinguishable particles with $H = \sum_{i=1}^{N} \frac{\vec{p_i}^2}{2m}$.

The partition function is:

$$Z(N,T,V) = \frac{1}{h^{3N}N!} \int \prod_{i} \mathrm{d}^{3}\vec{q_{i}} \mathrm{d}^{3}\vec{p_{i}} e^{-\beta \sum_{i=1}^{N} \frac{\vec{p_{i}}^{2}}{2m}} = \frac{1}{N!} \left(\frac{V}{\Lambda^{3}}\right)^{N} , \qquad (3.70)$$

where $\Lambda = \sqrt{\frac{h^2}{2\pi m k_B T}}$ is the thermal de Broglie wavelength. As we will see in the Chapter on statistical quantum mechanics, it is the length scale below which quantum effects start to play an important role and classical statistical mechanics is not sufficient anymore.

The Helmholtz free energy reads:

$$F = -k_B T \ln Z \simeq -Nk_B T \ln \left(\frac{eV}{N\Lambda^3}\right)$$
(3.71)

Notice that the computations for Z and F are much easier than the corresponding computations for Ω and S in the microcanonical ensemble.

Two-level systems: Consider N indistinguishable particles with energy $\varepsilon_i = 0, \varepsilon$. One microscopic configuration is $\mathcal{C} = \{\varepsilon_1, \ldots, \varepsilon_N\}$. The partition function and the Helmholtz free energy read:

$$Z = \sum_{\mathcal{C}} e^{-\beta \sum_{i} \varepsilon_{i}} = \sum_{\varepsilon_{1}} e^{-\beta \varepsilon_{1}} \cdot \sum_{\varepsilon_{2}} e^{-\beta \varepsilon_{2}} \cdots \sum_{\varepsilon_{N}} e^{-\beta \varepsilon_{N}} = \left(1 + e^{-\beta \varepsilon}\right)^{N}$$
(3.72)

$$F = -k_B T \ln Z = -Nk_B T \ln \left(1 + e^{-\beta\varepsilon}\right).$$
(3.73)

3.2.3 Fluctuations of energy and ensemble equivalence

Contrary to the microcanonical ensemble, the energy of configurations can now vary. Let's characterize the corresponding fluctuations.

Finite system: The partition function Eq. (3.58) closely resembles the moment-generating function for the energy:

$$\frac{\partial Z}{\partial(-\beta)} = \sum_{\mathcal{C}} E(\mathcal{C}) e^{-\beta E(\mathcal{C})} = Z \langle E \rangle \tag{3.74}$$

$$\Rightarrow \langle E^n \rangle = \frac{(-1)^n}{Z(\beta)} \frac{\partial^n Z(\beta)}{\partial \beta^n} \tag{3.75}$$

The relationship between the moment-generating function $Q(\lambda)$ and Z can be explicitly written by using the definition of $Q(\lambda)$:

$$Q(\lambda) = \langle e^{\lambda E} \rangle = \frac{1}{Z} \sum_{\mathcal{C}} e^{(\lambda - \beta)E} = \frac{Z(\beta - \lambda)}{Z(\beta)}.$$
(3.76)

Similarly, the cumulant-generating function is defined as:

$$\Psi(\lambda) = \ln Q(\lambda) = \ln Z(\beta - \lambda) - \ln Z(\beta) \quad \text{so that} \quad \langle E^n \rangle_c = (-1)^n \frac{\partial^n}{\partial \beta^n} (-\beta F) . \tag{3.77}$$

where $\langle E^n \rangle_c$ are the cumulants of the energy distribution. The Helmholtz free energy is (almost) the cumulant-generating function of the energy. As you can see, equilibrium statistical mechanics is (almost) a beautiful theory, albeit with some inconvenient factors and signs inherited from thermodynamics.

Probability distribution in the large size limit: Let us further characterize P(E). Using Eq. (3.43) for the case of system+thermostat one gets:

$$P(E) \propto \exp\left[-\frac{(E-E^{*})^{2}}{2k_{B}T^{2}\frac{c_{V}^{\text{th}}c_{V}}{c_{V}^{\text{th}}+c_{V}}}\right]$$
(3.78)

which, in the limit $c_V^{\text{th}} \gg c_V$, becomes

$$P(E) \propto e^{-\frac{(E-E^*)^2}{2k_B T^2 c_V}}$$
(3.79)

with E^* such that $\frac{\partial S_m}{\partial E}\Big|_{E^*} = \frac{1}{T}$. The same result can be derived directly from the canonical distribution $P(\mathcal{C}) = \frac{1}{Z} e^{-\beta E}$:

$$P(E) = \frac{1}{Z} \sum_{C|E(C)=E} e^{-\beta E} = \frac{\Omega(E)e^{-\beta E}}{Z} = \frac{e^{-\beta[E-TS]}}{Z}$$
(3.80)

$$Z = \sum_{\mathcal{C}} e^{-\beta E} = \sum_{E} \Omega(E) e^{-\beta E} = \frac{1}{\mathrm{d}E} \sum_{E} \mathrm{d}E \Omega(E) e^{-\beta E} \simeq \frac{1}{\mathrm{d}E} \int \mathrm{d}E e^{-\beta[E-TS(E)]} \quad (3.81)$$

E and S(E) are extensive quantities which scale $\sim N$ in the large N limit. As a consequence, for large system size, the integral is dominated by the maximum E^* such that $\frac{\partial S_m}{\partial E}\Big|_{E^*} = \frac{1}{T}$. As a consequence, $Z \approx e^{-\beta[E^* - TS_m(E^*)]}$ and $P(E) \simeq e^{-\beta[E - E^* - TS_m(E) + TS_m(E^*)]}$. Expanding around E^* and normalizing the resulting Gaussian distribution, one finally obtains:

$$P(E) = \frac{e^{\frac{-(E-E^*)^2}{2c_V k_B T^2}}}{\sqrt{2\pi c_V k_B T^2}}.$$
(3.82)

Ensemble equivalence: Since P(E) is Gaussian, $E^* = \langle E \rangle$. One can check that³

$$\lim_{\sigma \to 0} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-x_0)^2}{2\sigma^2}} = \delta(x-x_0)$$
(3.83)

so that P(E) converges in distribution to $\delta(E-E^*)$, which is the microcanonical distribution at an energy E^* such that

$$\left. \frac{\partial S_m}{\partial E} \right|_{E^*} = \frac{1}{T} \,. \tag{3.84}$$

In practice, in the microcanonical ensemble one fixes E and T adapts to satisfy Eq. (3.84). Conversely, in the canonical ensemble one fixes T and $\langle E \rangle = E^*$ adapts according to Eq. (3.84). As a consequence, for any observable O(E) that increases at most polynomially in E, in the large N limit:

$$\langle O(E) \rangle_{\text{micro}} = \int dE \delta(E - E^*) O(E) = O(E^*)$$
 (3.85)

$$\langle O(E) \rangle_{\text{cano}} = \int dE \omega(E) O(E) \frac{e^{-\beta E}}{Z} \underset{N \to \infty}{\sim} \int dE O(E) \delta(E - E^*) = O(E^*).$$
 (3.86)

³Simply integrate the left hand side against a function f(x) that you Taylor expand around x_0 . Carrying out the integrals order by order and sending σ to 0 will yield $f(x_0)$.



Figure 3.9: Left: Sketch of a scenario where S(E) is a concave function and the equivalence of ensemble holds. For any value of E, we can find a temperature such that $E^*(T) = E$. Right: Sketch showing a non-concave entropy function S(E) and its implications for the canonical ensemble. For a given value of T, there might be several value E_i^* such that $\frac{1}{T} = S'(E_i^*)$. In such a case, only the value of E_i^* that maximizes globally $S(E) - \frac{E}{T}$ is observed. In this case, E_3^* .

This means that, in the large-limit size, the microcanonical ensemble at energy E^* and the canonical ensemble at temperature T are equivalent, in the sense that they give the same predictions for the result of measurements, provided that E^* and T are linked via Eq. (3.84). This is the main result of the saddle-point approximation, valid for large N. It also explains why computing thermodynamic potentials in microcanonical and canonical ensemble will lead to equivalent forms in the thermodynamic limit, provided intensive and extensive control parameters like T and E are properly connected.

Ensemble inequivalence: an interesting exception. Equation (3.58) defines the set of accessible temperatures, and one can ask whether it is always possible to find a temperature such that $\langle E \rangle_{\text{cano}} = E^*$, for all E^* . In fact, this holds true only when the entropy function S(E) is concave, as sketched in the left panel of Fig. 3.9, ensuring that T(E) is a one-to-one mapping. Certain systems, such as those with long-range interactions, exhibit non-concave entropies, as sketched in the right panel of Fig. 3.9.

In such cases, the non-concave region in the microcanonical ensemble is not accessible in the canonical ensemble, leading to a first order phase transition, as sketched in Fig. 3.10. This demonstrates a fundamental inequivalence between ensembles. For more details, we refer the reader to [1].



Figure 3.10: In the presence of a non-concave entropy, the energy values inside the non-concave region $[E_1, E_2]$ cannot be accessed in the canonical ensemble. Changing T thus leads to to a discontinuous jump of energy at some temperature T_c which is the common tangent to $S_m(E)$ in E_1 and E_2 .



Figure 3.11: Schematic representation of a system in contact with a thermostat. The volume of the system is controlled by a piston that exerts a force F_w on the system.

Relation between F(T) and S(E): Using the result from the saddle-point approximation in Eq. (3.59), we get for the Helmholtz free energy:

$$F(T) = E^* - TS_m(E^*) \tag{3.87}$$

with $E^* = \operatorname{argmax} [E - TS_m(E)]$. Eq. (3.87) corresponds to the thermodynamic definition of Helmholtz free energy F = U - TS, where $U = \langle E \rangle = E^*$. Here, we use these terms interchangeably.

Notice that $-\beta F = \sup_{E} \left[k_B^{-1} S_m(E) - \beta E \right]$ is the Legendre transform of $\left(-k_B^{-1} S_m(E) \right)$ with respect to E. As before, the sign and factor of β is an unfortunate heritage from thermodynamics. If $S_m(E)$ is a concave function, we have complete equivalence of ensemble and one can obtain $S_m(E)$ from $F(\beta)$ by inverting the Legendre transform, $-k_B^{-1}S_m(E) = \sup_{\beta} \left[\beta F(\beta) - \beta E\right]$. To read more on ensemble equivalence, it is useful to rephrase equilibrium statistical mechanics, as done in the review article [2].

Gibbs entropy: So far, we have exclusively used the Boltzmann definition of entropy $S(E) = S_m(E) = k_B \ln \Omega(E)$. However, we can also define an entropy in the canonical ensemble, using Gibbs definition:

$$S_{c}(T) = -k_{B} \sum_{\mathcal{C}} P_{c}(\mathcal{C}) \ln P_{c}(\mathcal{C}) = -\frac{k_{B}}{Z} \sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})} \left[-\beta E(\mathcal{C}) - \ln Z\right]$$
(3.88)
$$= \frac{1}{T} \langle E(\mathcal{C}) \rangle + k_{B} \ln Z \underbrace{\sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})}}_{=1}$$
$$\Rightarrow S_{c}(T) = \frac{\langle E \rangle}{T} - \frac{F}{T}.$$
(3.89)

This function, $S_c(T)$, is, in general, different from $S_m(E) = k_B \log \Omega(E)$. However, in the large N limit, thanks to the saddle-point relation leading to $\langle E \rangle = E^*$ and $F = E^* - TS_m$, one recovers the equivalence of the two definitions of entropy: $S_c(T) = S_m(E^*(T))$. More generally, if microcanonical and canonical ensembles are equivalent at large N, we should be able to construct the same thermodynamics from each of them.

3.2.4 Thermodynamics from the canonical ensemble perspective

Let us now show that the mechanical pressure exerted by the system is $P = -\frac{\partial F}{\partial V}|_T$. Consider the setup sketched in the figure 3.11. The system exchanges energy with the thermostat and experiences a force F_w due to the piston. The piston wall can be modeled as a repulsive potential that diverges at position L, with $V_w(x) = \tilde{V}_w(x - L)$. The total Hamiltonian of the system is given by:

$$H_{\text{tot}} = H_0 + \sum_{i=1}^{N} \tilde{V}_w(x_i - L)$$
(3.90)

The total force exerted by the piston on the particles is $F_w = -\sum_{i=1}^N \partial_x V(x_i - L)$ and the variations of the volume due to the moving piston obey dV = AdL so that:

$$-\frac{\partial F}{\partial V}\Big|_{T} = -\frac{1}{A} \left.\frac{\partial F}{\partial L}\right|_{T} = -\frac{k_{B}T}{A} \left.\frac{\partial \ln Z}{\partial L}\right|_{T} = -\frac{k_{B}T}{AZ} \left.\frac{\partial Z}{\partial L}\right|_{T}$$
(3.91)

Using Eq. (3.90) in the definition of Z, one gets:

$$\frac{1}{Z}\frac{\partial Z}{\partial L} = \frac{1}{h^{3N}N!Z} \int \prod_{i} \mathrm{d}\Gamma_{i} \left(\sum_{i=1}^{N} \beta \partial_{x} V(x_{i} - L)\right) e^{-\beta H_{\mathrm{tot}}} = \beta \left\langle \sum_{i=1^{N}} \partial_{x} V(x_{i} - L)\right\rangle = \beta \langle F_{x} \rangle \qquad (3.92)$$

from which it follows that:

$$P = \frac{\langle F_x \rangle}{A} = -\left. \frac{\partial F}{\partial V} \right|_T \tag{3.93}$$

Consistency with the microcanonical result. In the microcanonical ensemble, we have shown that $dS = \frac{1}{T}dU + \frac{p}{T}dV$, where U = E, the imposed energy. We now have that F = U - TS, with $U = E^*(V,T)$. Differentiating F with respect to the volume thus yields

$$\frac{\partial F}{\partial V}\Big|_{T} = \left.\frac{\partial U}{\partial V}\right|_{T} - T \left.\frac{\partial S}{\partial V}\right|_{T} \,. \tag{3.94}$$

Since the first law states that $dS = \frac{dU}{T} + \frac{p}{T}dV$, we can express U as a function of T and V to get

$$dS = \frac{1}{T} \left[\left. \frac{\partial U}{\partial V} \right|_T dV + \left. \frac{\partial U}{\partial T} \right|_V dT \right] + \frac{p}{T} dV \quad \text{so that} \quad T \left. \frac{\partial S}{\partial V} \right|_T = P + \left. \frac{\partial U}{\partial V} \right|_T , \tag{3.95}$$

which yields back $P = -\frac{\partial F}{\partial V}|_T$, as expected.

Canonical entropy $S_c = -\frac{\partial F}{\partial T}\Big|_V$. Now that we have related $\frac{\partial F}{\partial V}\Big|_T$ to a thermodynamic observable, let us turn to $\frac{\partial F}{\partial T}\Big|_V$. Starting from the definition of F, we find:

$$\frac{\partial F}{\partial T}\Big|_{V} = \frac{\partial}{\partial T} \left(-k_{B}T \ln Z\right) = -\frac{k_{B}T \ln Z}{T} + \frac{1}{T} \underbrace{\frac{\partial}{\partial \beta} \ln Z}_{-\langle E \rangle} = \frac{F}{T} - \frac{\langle E \rangle}{T}$$
(3.96)

where we have used the fact that $\frac{\partial}{\partial T} = -\frac{1}{k_B T^2} \frac{\partial}{\partial \beta}$. It follows that:

$$\left. \frac{\partial F}{\partial T} \right|_{V} = -S_{c}(T, V) \underset{N \to \infty}{=} -S_{m}(U, V).$$
(3.97)

This is, again, consistent with the first law (check as an exercise, starting from the differential $dF = \frac{\partial F}{\partial T}\Big|_V dT + \frac{\partial F}{\partial V}\Big|_T dV = -SdT - pdV$).

Heat. Finally, we note that our system can now exchange energy with the environment in a form that is not work. This allows us to introduce the concept of heat δQ , which we define as "the energy exchanged with the environment that is not work." When the system is continuously transformed from one state to another, remaining equilibrated at all time (quasistatic, reversible transformation), then we have that

$$\delta Q = T \mathrm{d}S \;. \tag{3.98}$$

This is consistent with the relation $C_V = T \frac{\mathrm{d}S}{\mathrm{d}T} \simeq \frac{\delta Q}{\delta T}$ that we have used several times.

Summary. In the large N limit, the concentration of the measures in the statistical ensembles leads to ensemble equivalence (when the thermodynamic potentials is either strictly concave or strictly convex). The saddle point induces a one-to-one mapping between the control parameters, respectively T and Ein the canonical and microcanonical ensembles. As a consequence, we can choose our favorite ensemble to compute thermodynamic functions, they will yield the correct thermodynamic observables in the large system size limit.

Check for the ideal gas: We can check that the microcanonical and the canonical ensembles yield the same expressions for the physical quantities in the thermodynamic limit. For the microcanonical ensemble, we have computed:

$$S_m(E,V,N) = Nk_B \ln\left[\frac{Ve}{N} \left(\frac{4\pi meE}{3Nh^2}\right)^{3/2}\right]$$
(3.99)

to be compared with:

$$S_c(T, V, N) = -\frac{\partial F}{\partial T} = Nk_B \ln\left[\frac{Ve}{N}\left(\frac{2\pi mk_B T}{h^2}\right)\right]$$
(3.100)

Using that $S'_m(E) = \frac{3}{2}Nk_B\frac{1}{E} = \frac{1}{T}$, we find that $E^* = \frac{3}{2}Nk_BT$ so that $S_m(E^*, V, N) = S_c(T, V, N)$ and the two expressions coincide.

3.3 The grand canonical ensemble



Figure 3.12: Left: System exchanging energy (ΔE) and particles (ΔN) with a thermostat. Right: System exchanging energy and particles with a reservoir, which is itself in thermal equilibrium with a larger thermostat.

Let us now consider the case of 'open' systems, which can exchange particles and energy with a reservoir so that their particle number fluctuates. As we show below, when the total system is in the microcanonical ensemble, these systems will be endowed with a steady-state distribution that we can construct explicitly, which is called the grand canonical distribution.

To construct this distribution, we can proceed in two equivalent ways, as shown in Fig. 3.12. In the left panel, one starts with a scenario similar to what we did to construct the canonical distribution: a subsystem S exchange particles and energy with a thermostat and the total system is in the microcanonical ensemble. Alternatively, we can consider a system S that exchange particles with a reservoir, which itself exchanges energy with a larger thermostat, as illustrated on the right panel of Fig. 3.12.

Here, we follow this second route, which is less similar to what we have already done. Since the system + reservoir can exchange energy with the reservoir, they are in the canonical ensemble. We can thus describe the probability of the microstate $\{C_s, C_R\}$ by the Boltzmann weight:

$$P(\mathcal{C}_s, \mathcal{C}_{\rm res}) = \frac{1}{Z(T, N)} e^{-\beta [E(\mathcal{C}_s) + E(\mathcal{C}_{\rm res})]} , \qquad (3.101)$$

where β is the inverse temperature of the thermostat.

As before, a microstate for the system C_S defines a macrostate for the system + reservoir. By marginalizing on the reservoir, one obtains:

$$P(\mathcal{C}_S) = \sum_{\mathcal{C}_{\rm res}|N_S+N_{\rm res}=N} \frac{1}{Z_{\rm tot}(T,N)} e^{-\beta E(\mathcal{C}_{\rm res})-\beta E(\mathcal{C}_S)} = e^{-\beta E(\mathcal{C}_S)} \frac{1}{Z_{\rm tot}(T,N)} \underbrace{\sum_{\mathcal{C}_{\rm res}|N_{\rm res}=N-N_S} e^{-\beta E(\mathcal{C}_{\rm res})}}_{Z_{\rm res}(T,N-N_s)}.$$
(3.102)

Using the definition of Helmholtz free energy and expanding $F_{\rm res}(T, V_{\rm res}, N - N_S) \simeq F_{\rm res}(T, V_{\rm res}, N) - N_s \frac{F_{\rm res}}{\partial N}$, we obtain:

$$P(\mathcal{C}_S) = \frac{e^{-\beta E(\mathcal{C}_S) + \beta \mu N(\mathcal{C}_S)}}{Q}$$
(3.103)

where we defined the chemical potential of the reservoir:

$$\mu_{\rm res} = \frac{\partial F_{\rm res}(T, V, N)}{\partial N} \tag{3.104}$$

and the grand canonical partition function:

$$Q = \sum_{C_S} e^{-\beta E(C_S) + \beta \mu N(C_S)} = \sum_N e^{\beta \mu N} \sum_{C_S | N(C_S) = N} e^{-\beta E(C_S)} = \sum_N e^{\beta \mu N} Z_c(T, V, N) = \sum_{N, E} e^{\beta [\mu N - E + TS_m]} .$$
(3.105)

We also introduce the fugacity

$$z \equiv e^{\beta \mu} \tag{3.106}$$

so that

$$Q = \sum_{N} z^{N} Z_{c}(T, V, N) , \qquad (3.107)$$

and the grand potential:

$$G = -k_B T \ln Q \tag{3.108}$$

3.3.1 Fluctuations and large V limit

The quantity N now fluctuates and is determined, in the large V limit, by the chemical potential in Eq. (3.104). Notice that $\mu = \partial_N F \underset{N \to \infty}{\sim} \mathcal{O}(1)$ since $F \sim \mathcal{O}(N)$ is extensive. The chemical potential is an intensive quantity, like the temperature. To take the large system limit, we can now only send $V \to \infty$, as it is the sole remaining extensive parameter, keeping T and μ constant.

Fluctuations of N: Using $\frac{\partial}{\partial \mu} = \frac{\partial z}{\partial \mu} \frac{\partial}{\partial z} = \beta z \frac{\partial}{\partial z}$ we obtain the moment of N as:

$$\langle N^n \rangle = \frac{1}{Q} \frac{\partial^n Q}{\partial (\beta \mu)^n} = \frac{1}{Q} \left(z \frac{\partial}{\partial z} \right)^n Q .$$
(3.109)

The first derivative $\frac{\partial}{\beta\mu}$ means that we take a derivative of Q with respect to the variable $y = \beta\mu$, keeping β constant. A useful case is n = 1:

$$\langle N \rangle = \frac{\partial \ln Q}{\partial \beta \mu} = z \frac{\partial}{\partial z} \ln Q .$$
 (3.110)

The cumulants of N are obtained in a similar way:

$$\langle N^n \rangle = -\beta \left(z \frac{\partial}{\partial z} \right)^n G,$$
 (3.111)

where we have used that $\langle e^{\lambda N} \rangle = \frac{Q(\mu + \frac{\lambda}{\beta})}{Q(\mu)}$ to get the cumulant generating function $\psi(\lambda) = \ln Q(\mu + \frac{\lambda}{\beta}) - \ln Q$.

The typical fluctuations of N are then given by:

$$\langle N \rangle = -\frac{\partial}{\partial \mu} G \bigg|_{T}$$
(3.112)

$$\langle N^2 \rangle_c = \left. -\frac{1}{\beta} \left. \frac{\partial^2}{\partial \mu^2} G \right|_T = k_B T \frac{\partial}{\partial \mu} \langle N \rangle \,.$$
 (3.113)

Thus, the typical fluctuations scale as $\sqrt{\langle N^2 \rangle_c} \propto \sqrt{\langle N \rangle} \ll \langle N \rangle$. As with the relative fluctuations of E in the canonical ensemble, the relative fluctuations of N become negligible as $V \to \infty$.

3.3.2 Large V limit and thermodynamics

Starting from the definition of Q, Eq. (3.105), the grand canonical partition function can be approximated as:

$$Q = \sum_{N,E} e^{\beta[\mu N - E + TS_m]} \underset{V \to \infty}{\approx} e^{\beta \mu N^* - \beta E^* + \beta TS_m(E^*, N^*, V)}$$
(3.114)

where N^* and E^* maximize

$$\Psi_{GC}(N, E, V) = \mu N - E + TS_m(E, N, V).$$
(3.115)

This leads to the following conditions:

$$\frac{\partial \Psi}{\partial E} = 0 \Rightarrow \left. \frac{\partial S_m(E, N, V)}{\partial E} \right|_{E^*, N^*} = \frac{1}{T}$$
(3.116)

$$\frac{\partial \Psi}{\partial N} = 0 \Rightarrow \left. \frac{\partial S_m(E, N, V)}{\partial N} \right|_{E^*, N^*} = -\frac{\mu}{T}$$
(3.117)

59

This leads to

$$G(\mu, T, V) = E^*(\mu, T, V) - TS_m[(E^*(\mu, T, V), N^*(\mu, T, V), V] - \mu N^*(\mu, T, V)$$
(3.118)

Equivalently, one can also compute Q as:

$$Q = \sum_{N} e^{\beta \mu N} Z_c(N, V; T) = \sum_{N} e^{\beta \mu N - \beta F(N, V, T)}$$
(3.119)

and, using the saddle-point approximation, we find

$$\mu = \left. \frac{\partial F}{\partial N} \right|_{N^*, V, T} \tag{3.120}$$

with $N^* = \underset{N}{\operatorname{argmax}} [\mu N - F(N, V, T)]$. This then leads to:

$$G(\mu, T, V) = F(N^*(\mu, T, V), T, V) - \mu N^*(\mu, T, V) .$$
(3.121)

The consistency between the two computations above can be checked in the large V limit, starting from the expression for the Helmholtz free energy:

$$F(N,T,V) = E^* - TS_m(E^*, N, V)$$
(3.122)

$$= E^{*}(N, V, T) - TS_{m}(E^{*}(N, V, T), N, V)$$
(3.123)

where $\frac{\partial S_m}{\partial E}\Big|_{E^*,N} = \frac{1}{T}$. This immediately shows the consistency between Eq. (3.118) and Eq. (3.121). The consistency between Eq. (3.117) and (3.120) the stems from:

$$\frac{\partial F}{\partial N}\Big|_{T,V} = \frac{\partial E^*}{\partial N}\Big|_{T,V} - T \left.\frac{\partial S_m}{\partial N}\Big|_{E^*,V} - T \underbrace{\frac{\partial S_m}{\partial E}\Big|_{N,V}}_{1/T} \frac{\partial E^*}{\partial N}\Big|_{T,V} = -T \left.\frac{\partial S_m}{\partial N}\Big|_{E^*,N} \tag{3.124}$$

So far, we have ignored the variations of F and S with respect to N. We can now introduce them back:

$$dF = \frac{dF}{dV}dV + \frac{dF}{dT}dT + \frac{dF}{dN}dN = \frac{p}{T}dV - \frac{1}{T}dE - \frac{\mu}{T}dN$$
(3.125)

$$dS = \frac{dS}{dV}dV + \frac{dS}{dE}dE + \frac{dS}{dN}dN = \frac{p}{T}dV - \frac{1}{T}dE - \frac{\mu}{T}dN$$
(3.126)

Notice that Eq. (3.126) is the first law of thermodynamics. Then, for the grand canonical potential $G = E - TS - \mu N = F - \mu N$ leads to:

$$dG = -pdV - SdT - Nd\mu \tag{3.127}$$

from which one obtains alternative expressions for the entropy $S = -\frac{\partial G}{\partial T}\Big|_{V,\mu}$ and the pressure $p = -\frac{\partial G}{\partial V}\Big|_{T,\mu}$. All these relations are consistent with the microcanonical and canonical definitions in the large V limit.

3.3.3 Ideal gas

The grand partition function for the ideal gas is:

$$Q = \sum_{N=0}^{\infty} e^{\beta\mu N} Z_c(N, T, V) = \sum_N e^{\beta\mu N} \frac{1}{N!} \left(\frac{V}{\Lambda^3}\right)^N = e^{Vz/\Lambda^3}$$
(3.128)

It follows that, in the grand canonical ensemble, N is distributed according to a Poisson distribution:

$$P_{GC}(N) = \frac{1}{N!} \left(\frac{Vz}{\Lambda^3}\right)^N e^{-\frac{Vz}{\Lambda^3}}$$
(3.129)

with Poisson parameter $\langle N \rangle = \frac{Vz}{\Lambda^3}$.

For the ideal gas, the fugacity z is thus directly related to the average density $\rho_0 = \frac{\langle N \rangle}{V}$ in the system:

$$z = \frac{\langle N \rangle}{V} \Lambda^3 = \rho_0 \Lambda^3 , \qquad (3.130)$$

and the chemical potential can then also be expressed as a function of the average density :

$$\mu = k_B T \ln \frac{\Lambda^3 \langle N \rangle}{V} = k_B T \ln(\rho_0 \Lambda^3)$$
(3.131)

Finally, the grand potential and the pressure are:

$$G = -k_B T \frac{Vz}{\Lambda^3} \tag{3.132}$$

$$P = k_B T \frac{\langle N \rangle}{V} \tag{3.133}$$

where in the second line one recognizes the ideal gas law.

3.4 Thermodynamics

In the large-size limits, we have seen that many systems admits ensemble equivalence. Provided the right relations between thermodynamic variables are enforced, what is computed in one ensemble can be used to predict the value of observables in other ensembles. This is very powerful. For instance, it is relatively easy to compute the thermodynamic properties of a Boson gas in the grand-canonical ensemble, but experiments are better described in the canonical ensemble. We can thus use ensemble equivalence to compute observables in the ensemble where theory is easiest and export the results to ensemble that are relevant experimentally.

3.4.1 Thermodynamics variables

We distinguish extensive thermodynamic variables N, V, E and intensive thermodynamic variables T, p, μ . When building ensembles, we can choose to fix either the extensive or the intensive variables. This gives $2^3 - 1 = 7$ different ensembles with at least one extensive variable. All those ensembles have an associated thermodynamic potential. In this course, we have seen

$$(E, V, N) \rightarrow S_m$$
 (3.134)

$$(T, V, N) \rightarrow F$$
 (3.135)

$$(T, V, \mu) \rightarrow G$$
 (3.136)

In the large size limit, all ensembles lead to consistent thermodynamic relations provided the independent variables are related by the saddle point relations, for e.g. $\frac{1}{T} = \frac{\partial S}{\partial E}|_{U}$, provided the thermodynamic potentials have the right convexity properties. In turn, the saddle point relations lead to Legendre transforms relations between the thermodynamic potential (e.g. $\beta F = \beta U - S/k_B$).

In statistical mechanics, because we know which variable is fixed in which ensemble, there is no confusion about what is a function of what. In the large size limit, when all ensembles becomes equivalent, confusion steps in because thermodynamics observables are not independent. This requires either extremely heavy notations, as in Eq. (3.118), or some extra care.

The evil crime of thermodynamics: In the thermodynamic limit, each observable can be expressed in terms of different variables. For instance, we have

$$S_{GC}(T, V, \mu) = S_C(T, V, N^*(T, V, \mu)) = S_m(E^*(T, V, N^*(T, V, \mu)), V, N^*(T, V, \mu)).$$
(3.137)

The three functions S_{GC} , S_C and S_m are distinct functions, with distinct set of arguments. When their arguments are related through the saddle point relations that we derived earlier, they yield the same value. Because the notation (3.138) is horribly heavy, it is tempting to drop the subscript and write all these functions as S. Further, it is assumed that the arguments are not independent. This lightens a lot our notations since we can now write:

$$S(T, V, \mu) = S(T, V, N) = S(E, V, N), \qquad (3.138)$$

even though one has to remember that these are different mathematical functions and that N and E are function of T, V, and μ in the equation above.

The standard chain rule: This notation becomes a huge source of errors when one whishes to take derivatives.

If one wants to take a derivative of S with respect to V, we have to be clear on which of these functions we are considering. For instance, the chain rule tells us that:

$$\partial_V S_{GC}(T, V, \mu) = \partial_V S_C(T, V, N^*(T, V, \mu)) + \partial_N S_C(T, V, N^*(T, V, \mu)) \partial_V N^*(T, V, \mu) .$$
(3.139)

This is standard mathematics and there are no possible mistakes.

However, if we drop the subscripts, how can we know which S are we considering? The notational solution is to indicate what are the variables that are being kept fixed: here, keeping T and μ fixed when taking a derivative with respect to V refers to S_{GC} , while keeping T and N fixed refers to S_C . So that one rewrites Eq. (3.138) as:

$$\frac{\partial S}{\partial V}\Big|_{T,\mu} = \frac{\partial S}{\partial V}\Big|_{T,N} + \frac{\partial S}{\partial N}\Big|_{T,V} \frac{\partial N}{\partial V}\Big|_{T,\mu}$$
(3.140)

The standard but heavy maths of Eq. (3.139) have now become lighter but much less transparent... To specify which function is being used when taking a partial derivative, thermodynamics thus employs the notation $\frac{\partial S_m}{\partial X}\Big|_{Y,Z}$, or $\frac{\partial S_m}{\partial X}\Big|_{Y,Z}$, which indicates that the considered function is S(X, Y, Z).

Another example of the standard chain rule is to differentiate $S_c(T, V, N) = S_m(E(T, V, N), N, V)$ with respect to the volume (dropping the star notation):

$$\frac{\partial S_c(T,V,N)}{\partial V} = \frac{\partial S_m(E(T,V,N),V,N)}{\partial V} + \frac{\partial S_m(E(T,V,N),V,N)}{\partial E} \cdot \frac{\partial E(T,V,N)}{\partial V}$$
(3.141)

3.4. THERMODYNAMICS

This simply becomes

$$\underbrace{\frac{\partial S}{\partial V}}_{N,T,V\Rightarrow S \text{ canonical}} = \underbrace{\frac{\partial S}{\partial V}}_{N,E,V\Rightarrow S \text{ microcanonical}} + \underbrace{\frac{\partial S}{\partial E}}_{N,V} \cdot \underbrace{\frac{\partial E}{\partial V}}_{N,T,V\Rightarrow E \text{ canonical}}$$
(3.142)

The modified chain rule: In addition to the standard chain rule above, the fact that thermodynamic observables are dependent also lead to more elaborate relationships between partial derivatives.

A relation such as F = U - TS expresses an interdependence among the potentials F, U, S. One may express this as the constraint g(U, S, F) = 0, whose differential is:

$$dg = 0 = \frac{\partial g}{\partial U} \Big|_{S,F} dU + \frac{\partial g}{\partial S} \Big|_{U,F} dS + \frac{\partial g}{\partial F} \Big|_{U,S} dF.$$
(3.143)

Notice that, when S, F are fixed, the function $g(U, S, F) = g_{SF}(U)$ becomes a 1D function of a sole variable, which can be inverted.



Figure 3.13: Here $\left.\frac{\partial g}{\partial u}\right|_{u^*} = \frac{\Delta g}{\Delta u}$ can be inverted as $\left(\left.\frac{\partial g}{\partial u}\right|_{u^*}\right)^{-1} = \frac{\Delta u}{\Delta g} = \left.\frac{\partial u}{\partial g}\right|_{g^*}$.

Multiplying Eq. (3.143) by $\left.\frac{\partial U}{\partial g}\right|_{F,S}$, one gets:

$$dU + \frac{\partial U}{\partial g}\Big|_{F,S} \frac{\partial g}{\partial F}\Big|_{U,S} dF + \frac{\partial U}{\partial g}\Big|_{F,S} \frac{\partial g}{\partial S}\Big|_{F,U} dS = 0$$
(3.144)

If follows that:

$$\frac{\partial U}{\partial S}\Big|_{F,T} = -\frac{\partial U}{\partial g}\Big|_{F,S} \frac{\partial g}{\partial F}\Big|_{U,S}$$
(3.145)

$$\frac{\partial S}{\partial F}\Big|_{U,T} = -\frac{\partial S}{\partial g}\Big|_{U,F} \frac{\partial g}{\partial F}\Big|_{U,S}$$
(3.146)

$$\frac{\partial F}{\partial U}\Big|_{S,T} = -\frac{\partial F}{\partial g}\Big|_{U,S} \frac{\partial g}{\partial U}\Big|_{S,F}$$
(3.147)

Multiplying the left-hand sides of Eqs.(3.145)-(3.147), we find:

$$\frac{\partial U}{\partial S}\Big|_{F,T} \left. \frac{\partial S}{\partial F} \right|_{U,T} \left. \frac{\partial F}{\partial U} \right|_{S,T} = -1.$$
(3.148)

Equivalently:

$$\frac{\partial S}{\partial U}\Big|_{F,T} = -\frac{\partial S}{\partial F}\Big|_{U,T} \frac{\partial F}{\partial U}\Big|_{S,T}.$$
(3.149)

Equations (3.148) and (3.149) are often called the modified chain rule. Indeed, Eq. (3.149) looks like a chain rule with a minus sign. This is due to the fact that, once T is fixed, S, U and F are not independent anymore so that the standard chain rule cannot be applied.

3.4.2 Thermodynamic relations

The modified and the standard chain rules are essential tools to derive thermodynamic relations.

Additionally, the first law of thermodynamics $dU = TdS - pdV + \mu dN$, provides a fundamental relationship between changes in energy/entropy and variations in the extensive variables. It can be used to change variables. For instance, to express U(T, V, N), we expand dS as:

$$dS = \frac{\partial S}{\partial T} \Big|_{N,V} dT + \frac{\partial S}{\partial N} \Big|_{V,T} dN + \frac{\partial S}{\partial V} \Big|_{N,T} dV.$$
(3.150)

Replacing this into the first law yields:

$$dU = T \left. \frac{\partial S}{\partial T} \right|_{N,V} dT + \left(T \left. \frac{\partial S}{\partial V} \right|_{N,T} - p \right) dV + \left(T \left. \frac{\partial S}{\partial N} \right|_{V,T} + \mu \right) dN \,. \tag{3.151}$$

From there, we can get, for instance

$$\left. \frac{\partial U}{\partial N} \right|_{T,V} = T \left. \frac{\partial S}{\partial N} \right|_{V,T} + \mu \tag{3.152}$$

Additional thermodynamic relations arise from the saddle-point and Legendre transform such as F = U - TS or $G = U - TS + \mu N$.

Another source of thermodynamic relations is to invoke the extensivity of thermodynamic potentials. For instance, the extensivity of the energy implies that:

$$E(\lambda S, \lambda V, \lambda N) = \lambda E(S, V, N)$$
(3.153)

Taking a derivative with respect to λ and setting $\lambda = 1$ we get:

$$S\underbrace{\frac{\partial E}{\partial S}\Big|_{N,V}}_{T} + V\underbrace{\frac{\partial E}{\partial V}\Big|_{S,N}}_{=n} + N\underbrace{\frac{\partial E}{\partial N}\Big|_{S,V}}_{=n} = E(S,V,N)$$
(3.154)

$$\Rightarrow E = ST - PV + \mu N \tag{3.155}$$

Similarly, $G(T, \lambda V, N) = \lambda G(T, V, N)$ and, taking a derivative with respect to λ leads to:

$$G(T, V, N) = V \underbrace{\frac{\partial G}{\partial V}}_{-p} = -pV$$
(3.156)

Notice that by taking the differential of Eq. (3.155) and using the first law one obtains the Gibbs Duham relation:

$$SdT - VdP + Nd\mu = 0 \tag{3.157}$$

which relates variations between intensive thermodynamic variables.

Finally, the Maxwell relations relate the second derivative of thermodynamic variables, making use of Schwarz's theorem:

$$\frac{\partial}{\partial A} \left(\left(\frac{\partial X}{\partial B} \right)_{A,C} \right)_{B,C} = \frac{\partial}{\partial B} \left(\left(\frac{\partial X}{\partial A} \right)_{B,C} \right)_{A,C}.$$
(3.158)

For instance, given $\frac{\partial E}{\partial S}|_{N,V} = T$ and $\frac{\partial E}{\partial N}|_{S,V} = \mu$, it follows:

$$\left. \frac{\partial T}{\partial N} \right|_{V,S} = \left. \frac{\partial \mu}{\partial S} \right|_{N,V} \,. \tag{3.159}$$

Application: Number fluctuations and compressibility

In the grand canonical ensemble, we have seen that $\langle N^2 \rangle_c = k_B T \left. \frac{\partial \langle N \rangle}{\partial \mu} \right|_{T,V}$. While this is a useful relation to predict that $\langle N^2 \rangle_c \propto N$, measuring the right-hand side is difficult experimentally. Can we thus re-express the right-hand side into something that would be easier to measure?

Let us start by using the modified chain rule. At fixed T, the variables N, V and μ are dependent, so that

$$\frac{\partial N}{\partial \mu}\Big|_{V,T} = -\left.\frac{\partial N}{\partial V}\right|_{\mu,T} \left.\frac{\partial V}{\partial \mu}\right|_{T,N} \tag{3.160}$$

Then, let's say that we replace $V(\mu, T, N)$ in the last equation by $V(P(\mu, T, N), T, N)$ and that we use extensivity to write $N = \rho_0 V$, leading to

$$\frac{\partial N}{\partial \mu}\Big|_{V,T} = -\rho_0 \left. \frac{\partial V}{\partial p} \right|_{T,N} \left. \frac{\partial p}{\partial \mu} \right|_{T,N} \tag{3.161}$$

Now, this is progress since $\frac{\partial V}{\partial p}\Big|_{T,N}$ is measurable in experiments (applying pressure and measuring volume). This coefficient is so useful, that it has a name: the isothermal compressibility

$$\kappa_T = -\frac{1}{V} \left. \frac{\partial V}{\partial p} \right|_{T,N} \,, \tag{3.162}$$

where the $\frac{1}{V}$ prefactor is here to make κ_T an intensive variable.

To compute $\frac{\partial p}{\partial \mu}\Big|_{T,N}$, we now use a Maxwell relation. To identify the right potential, we note that we have to think about an ensemble in which P is a function and μ a variable. We can use the grand-canonical ensemble, where $dG = -SdT - pdV - Nd\mu$ to get:

$$\left. \frac{\partial p}{\partial \mu} \right|_{T,V} = \left. \frac{\partial N}{\partial V} \right|_{T,\mu} = \rho_0 \,. \tag{3.163}$$

It follows that $\left. \frac{\partial N}{\partial \mu} \right|_{V,T} = \rho_0^2 V \kappa_T = \frac{N^2}{V} \kappa_T$. From this, we derive the fluctuation-dissipation relation:

$$\frac{\langle N^2 \rangle_c}{\langle N \rangle^2} = \frac{k_B T}{V} \kappa_T \,. \tag{3.164}$$

Bibliography

- [1] Alessandro Campa, Thierry Dauxois, and Stefano Ruffo. Statistical mechanics and dynamics of solvable models with long-range interactions. *Physics Reports*, 480(3-6):57–159, 2009.
- [2] Hugo Touchette. The large deviation approach to statistical mechanics. *Physics Reports*, 478(1-3): 1–69, 2009.