

Introduction to Statistical Mechanics

Salvatore Assenza

September 30, 2021

1 Introduction

Statistical mechanics was born with the ambitious goal of deriving the laws of thermodynamics from a statistical treatment of the microscopic degrees of freedom of a system. Even more strikingly, this was done at a time when the very concept of the atomistic picture of the world was highly questioned. As we will see, statistical mechanics is a natural framework to understand the laws of thermodynamics and in particular to have a clear interpretation of what is entropy and how it is related to irreversibility. The statistical approach is rooted in the observation that a typical thermodynamic system is made of a huge number of particles (at least of the order of Avogadro's number, i.e. $N \sim 10^{23}$). Far from phase transitions (such as ice melting into water), computing averages is typically enough. In contrast, close to phase transitions also fluctuations become important, and including them in the treatment is fundamental in order to properly retrieve the corresponding thermodynamic laws. Biophysical systems somehow borrow this last feature: since they can be relatively small, ($N \sim 10 - 100$), one can use statistical mechanics to properly describe them, although fluctuations are often important features to be considered.

2 Some useful mathematical tools

Before starting, it is important to summarize some concepts that will be used in this document.

2.1 Basic combinatorics

First, we introduce some results from combinatorics.

1. Imagine to have N elements that you want to order in some way. For instance, we want to create sequences with the figures 1, 2, 3 where each number appears exactly once. In this case, it is easy to enumerate them. All the possible sequences that we can form (called *permutations*) are 1 – 2 – 3, 1 – 3 – 2, 2 – 1 – 3, 2 – 3 – 1, 3 – 1 – 2, 3 – 2 – 1. Thus, the number of possible permutations of 3 elements is 6. Rather than listing all of them, we note that for the first number of the sequence we had 3 possible choices; for the second number, only 2 choices were available, since we had already used one of the figures; finally, for the last element of the sequence we had just 1 choice (i.e. no choice at all). Based on this, we can compute the number of permutations as $3 \cdot 2 \cdot 1 = 6$, which is the same result that we found by counting all the possible sequences. The advantage of this second approach is that it can be easily generalized to N elements: we have N choices for the first number, $N - 1$ for the second one, etc. Therefore, the total number P_N of permutations of N elements is

$$P_N = N \cdot (N - 1) \cdot (N - 2) \cdot \dots \cdot 2 \cdot 1 \equiv N! . \quad (1)$$

The operation $N!$ is called the *factorial* of N . Note that this problem is equivalent to the problem of assigning the positions to N distinguishable particles choosing from N possible spots, without the possibility of having two particles in the same spot.

2. We consider the same problem as in point 1, but this time we want to create a sequence of length $k \leq N$ (one speaks about *dispositions*). For instance, we decide to create sequences of length $k = 2$ with the figures 1, 2, 3, without allowing for repetition of each figure. It is easy to list the sequences: 1 – 2, 1 – 3, 2 – 1, 2 – 3, 3 – 1, 3 – 2. Again, we have 6 possibilities (in this case we have the same number as before, but it is just a special case). We note that we can follow the same argument as above, but we have to stop after k elements. Hence, we have N choices for the first element, $N - 1$ for the second, etc. until we arrive at $N - k + 1$ for the k th element. The total number of dispositions $D_{N,k}$ is thus $D_{N,k} = N \cdot (N - 1) \cdot (N - 2) \cdot \dots \cdot (N - k + 1)$. We note that these are the same numbers used to build $N!$ from which we cut out the terms $(N - k), (N - k - 1), \dots, 2, 1$. Hence, we can write

$$D_{N,k} = \frac{N \cdot \dots \cdot (N - k + 1) \cdot (N - k) \cdot (N - k - 1) \cdot \dots \cdot 2 \cdot 1}{(N - k) \cdot (N - k - 1) \cdot \dots \cdot 2 \cdot 1} = \frac{N!}{(N - k)!} . \quad (2)$$

Note that this problem is equivalent to the problem of assigning the positions to N distinguishable particles choosing from k possible spots, without the possibility of having two particles in the same spot.

3. We consider the same problem as in point 2, but this time we do not care about the order of the sequence. (one speaks about *combinations*). For instance, the sequence 3 – 1 and the sequence 1 – 3 are now considered equivalent, since the only important thing is the content of the sequence. Considering $N = 3$ and $k = 2$, we thus have as possible combinations 1 – 2, 1 – 3, 2 – 3, i.e. there are only 3 cases. The total number of combinations $C_{N,k}$ can be easily computed as the ratio between the number of dispositions $D_{N,k}$ and the number of possible permutations of k objects, P_k . We thus have

$$C_{N,k} = \frac{D_{N,k}}{P_k} = \frac{N!}{k!(N-k)!} . \quad (3)$$

Note that this problem is equivalent to the problem of assigning the positions to N indistinguishable particles choosing from k possible spots, without the possibility of having two particles in the same spot.

4. We consider the same problem as in point 1, but this time we allow for repetition of the elements. For instance, the possible sequences with 1, 2, 3 are now 1 – 1 – 1, 1 – 1 – 2, 1 – 1 – 3, 1 – 2 – 1, 1 – 2 – 2, 1 – 2 – 3, 1 – 3 – 1, 1 – 3 – 2, 1 – 3 – 3, 2 – 1 – 1, 2 – 1 – 2, 2 – 1 – 3, 2 – 2 – 1, 2 – 2 – 2, 2 – 2 – 3, 2 – 3 – 1, 2 – 3 – 2, 2 – 3 – 3, 3 – 1 – 1, 3 – 1 – 2, 3 – 1 – 3, 3 – 2 – 1, 3 – 2 – 2, 3 – 2 – 3, 3 – 3 – 1, 3 – 3 – 2, 3 – 3 – 3. There are a total of 27 possible sequences. Following the same approach as above, this time we have N choices for the first element, N choices for the second, etc. Thus, the number of possible sequences is

$$P_N^r = N^N , \quad (4)$$

where the r stands for “repetition”. Note that this problem is equivalent to the problem of assigning the positions to N distinguishable particles choosing from N possible spots, with the possibility of having two particles in the same spot. Analogously, we can consider the number $D_{N,k}^r$ of dispositions with repetition, finding

$$D_{N,k}^r = N^k . \quad (5)$$

2.2 Probability in a nutshell

We also have to list some key elements from probability theory.

1. If in a system there are \mathcal{N} possible equiprobable states, the probability of each of them occurring is simply $1/\mathcal{N}$. For instance, when throwing a dice there are six possible outcomes, so that $\mathcal{N} = 6$ and the probability of each of them occurring is $1/6$.

2. More in general, the probability for a given outcome (or *event* X) is given by the number of cases $g(X)$ compatible with it (called its *degeneracy*) divided by the total number of possibilities \mathcal{N} , i.e. $p(X) = g(X)/\mathcal{N}$. For instance, the probability that by throwing a dice we get a number larger than 4 is $2/6$.
3. Now, let us throw two dices. The state of the system is defined by the couple (s_1, s_2) , where $s_1, s_2 = 1, 2, \dots, 6$. Hence, there are in total $6 \times 6 = 36$ possible states. The probability for each state is $1/36$. For instance, the probability that the first dice gives 1 and the second dice gives 4 is equal $1/36$. Note that $1/36 = 1/6 \times 1/6$. This is not a coincidence: *the probability of the co-occurrence of two independent events X_1, X_2 is equal to the product of the probabilities of each of them occurring, $p_{12}(X_1, X_2) = p_1(X_1)p_2(X_2)$* . Note that for this to be true the events have to be *independent*. For instance, in the previous example the event X_1 was “dice 1 gives 1” and the event X_2 was “dice 2 gives 4”, which are independent of each other. As a counterexample, we can consider the same event X_1 while event X_2 is “dice 1 gives a number lower than 5”. We note that X_2 occurs with probability $5/6$ and, if this happens, it automatically includes X_1 as a byproduct. Therefore, the probability of co-occurrence in this case is simply $5/6 \neq 1/6 \times 5/6$.
4. Considering again one dice, let us consider the two events “dice 1 gives 1” and the event X_2 was “dice 1 gives 4”. We ask what is the probability that either of them happens. Since there are two favourable cases out of six, we get $p(X_1 \text{ or } X_2) = 2/6 = p(X_1) + p(X_2)$. Again, this is not a coincidence: *disjoint events occur with an overall probability which is equal to the sum of the probabilities of each event*. The keyword to remember here is *disjoint*, meaning that there is no overlap between the two events. As a counterexample, let us consider as X_1 the event “dice 1 gives a number between 2 and 5, extremes included” and as X_2 the event “dice 1 gives a number above 3”. We see that $p(X_1) = 4/6$ and $p(X_2) = 3/6$. However, the event X_1 or X_2 corresponds to “dice 1 gives a number above 1”, which occurs with probability $p(X_1 \text{ or } X_2) = 5/6 \neq p(X_1) + p(X_2)$. We also notice *en passant* that $p(X_1) + p(X_2) > 1$, which does not make any sense! The problem with applying this formula is that the two events are overlapping, so that by summing the probabilities we count twice the overlapping part.
5. By construction, the probability p of any event is always contained in the interval $[0, 1]$. $p = 0$ means that the event is impossible (e.g. “dice 1 gives a number which is at the same time below 2 and above 4”), while $p = 1$ means that the event occurs deterministically (e.g. “dice 1 gives a positive number”).
6. In the case of continuous variables, one has to adapt the formalism. For instance, imagine to extract a random real number in the interval $[0, 1]$. By definition, the *probability density* $p(x)$ is such that $p(x)dx$ is the infinitesimal probability that the real number is found in the interval $[x, x + dx]$. The probability of an event is obtained by integrating $p(x)$ over a suitable interval. For instance, the probability that the number is found in the

interval $[0.2, 0.7]$ is $\int_{0.2}^{0.7} p(x)dx$. Note that by integrating $p(x)$ over the whole interval gives 1 by construction (*normalization condition*). For instance, in the current example $\int_0^1 p(x)dx = 1$. Note that any choice for $p(x)$ is valid, as long as $p(x) > 0$ for any x and the normalization condition is satisfied. As a simple example, let us assume that the random number is extracted from $[0, 1]$ with uniform probability, i.e. $p(x) = p$ independently from x . From the normalization condition, we get $1 = \int_0^1 p(x)dx = p \int_0^1 dx = p$, i.e. $p(x) = 1$ for any x . Hence, the probability that the random number lies in the interval $[0.5, 1]$ is $\int_{0.5}^1 p(x)dx = 1 - 0.5 = 0.5$. As a second example, let us instead assume that $p(x) = c \cdot x^2$ (which gives more weight to larger numbers). The constant c is found by imposing the normalization condition: $1 = \int_0^1 p(x)dx = c \int_0^1 x^2 dx = c/3 \Rightarrow c = 3$. Hence, $p(x) = 3x^2$. The probability that x lies in the interval $[0.5, 1]$ is now $\int_{0.5}^1 p(x)dx = (c/3)x^2|_{0.5}^1 = 1 - 0.5^3 = 0.875$. As expected, this probability is larger than 0.5 since more weight was given to larger numbers.

7. Let us consider a certain continuous variable with probability density $p(x)$. Now, let us consider a second variable $y = y(x)$. Since y depends on x , its value will also be random with a certain probability density $q(y)$. What is the relation between $p(x)$ and $q(y)$? To answer this question, remember that the probability the y lies in the range $[y_A, y_B]$ is $\int_{y_A}^{y_B} q(y)dy$. We limit our discussion to the case in which $y(x)$ can be inverted, so that each point y has only one corresponding value of x . By performing a change of variables in the integral, we get

$$\int_{y_A}^{y_B} q(y)dy = \int_{x_A}^{x_B} q(y(x)) \frac{dy}{dx} dx. \quad (6)$$

The function $y(x)$ is monotonous (otherwise it cannot be inverted). If $dy/dx > 0$, then $x_B > x_A$ (since $y_B > y_A$) and the probability that x lies in the range $[x_A, x_B]$ is $\int_{x_A}^{x_B} p(x)dx$. Comparing to the right-hand side of the previous equation and given the arbitrariness of the interval, we thus get $q(y) = p(x)/(dy/dx)$, i.e. we obtain q by renormalizing p by the derivative of $y(x)$. In the case in which $dy/dx < 0$, we have $x_B < x_A$, and the probability that x lies in the range $[x_B, x_A]$ is $\int_{x_B}^{x_A} p(x)dx = -\int_{x_A}^{x_B} p(x)dx$. Hence, we get $q(y) = -p(x)/(dy/dx)$. The two results can be joined together by writing $q(y) = p(x)/|dy/dx|$. Note that the denominator goes to zero at a stationary point, but for our applications the integral will not be affected. As a simple example, we consider for p the probability density associated to a real number extracted with uniform probability from the range $[0, 1]$. As we saw in the previous point, we thus have $p(x) = 1$. Now, we ask: what is the probability density of the cubic root of such a number? Hence, we are considering $y = x^{1/3} \Rightarrow dy/dx = 1/(3x^{2/3}) = 1/(3y^2) \Rightarrow q(y) = 3y^2$. This is the second case studied in the previous point (the function for the present example was chosen on purpose). Note that the integration interval does not have to be the same. For instance, if we choose $y(x) = 5x + 7$ the range of y values will be $[7, 12]$ and the probability density of y is $q(y) = 1/5$. As a further example, we now ask what is the probability

density of the variable $y = 5 - x^2$, which is a decreasing function of x . Indeed, in this case $dy/dx = -2x = -2\sqrt{5-y} < 0$. By applying the formula obtained above, we get $q(y) = 1/|-2\sqrt{5-y}| = 1/(2\sqrt{5-y})$.

2.3 Gaussian Integrals

Another important tool that we will need are the Gaussian integrals, i.e. integrals involving the Gaussian function e^{-ax^2} , where $a > 0$ is a constant. First of all, we start with the integral

$$G_1(a) \equiv \int_0^\infty x e^{-ax^2} dx = -\frac{e^{-ax^2}}{2a} \Big|_0^\infty = \frac{1}{2a}. \quad (7)$$

Now, we consider the integral

$$G_0(a) \equiv \int_0^\infty e^{-ax^2} dx. \quad (8)$$

In order to compute $G_0(a)$, we note that $\int_{-\infty}^\infty e^{-ax^2} dx = 2G_0(a)$, so that taking the square we find

$$4G_0(a)^2 = \int_{-\infty}^\infty e^{-ax^2} dx \int_{-\infty}^\infty e^{-ay^2} dy = \int_{-\infty}^\infty dx \int_{-\infty}^\infty dy e^{-a(x^2+y^2)}. \quad (9)$$

Considering polar coordinates (ρ, ϕ) , we have

$$x = \rho \cos \phi \quad y = \rho \sin \phi, \quad (10)$$

with $\rho > 0$ and $0 \leq \phi < 2\pi$. With the new coordinates, $x^2 + y^2 = \rho^2$ and $dx dy = \rho d\rho d\phi$. In the latter term, the prefactor ρ comes from the Jacobian of the set of transformations in Eq.(10). Substituting in Eq.(9), we obtain

$$4G_0(a)^2 = \int_0^\infty d\rho \int_0^{2\pi} d\phi \rho e^{-a\rho^2} = 2\pi G_1(a) = \frac{\pi}{a}. \quad (11)$$

Hence, we finally obtain

$$G_0(a) \equiv \int_0^\infty e^{-ax^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{a}}. \quad (12)$$

We further consider the integral

$$G_2(a) \equiv \int_0^\infty x^2 e^{-ax^2} dx. \quad (13)$$

We note that

$$\frac{dG_0(a)}{da} = \frac{d}{da} \int_0^\infty e^{-ax^2} dx = - \int_0^\infty x^2 e^{-ax^2} dx = -G_2(a). \quad (14)$$

Therefore, deriving Eq.(12) with respect to a we obtain

$$G_2(a) \equiv \int_0^\infty x^2 e^{-ax^2} dx = \frac{\sqrt{\pi}}{4a^{\frac{3}{2}}}. \quad (15)$$

Finally, we also consider

$$G_4(a) \equiv \int_0^\infty x^4 e^{-ax^2} dx. \quad (16)$$

We note that

$$\frac{dG_2(a)}{da} = \frac{d}{da} \int_0^\infty x^2 e^{-ax^2} dx = - \int_0^\infty x^4 e^{-ax^2} dx = -G_4(a). \quad (17)$$

Therefore, deriving Eq.(15) with respect to a we obtain

$$G_4(a) \equiv \int_0^\infty x^4 e^{-ax^2} dx = \frac{3\sqrt{\pi}}{8a^{\frac{5}{2}}}. \quad (18)$$

This procedure can be iterated to find a general formula for $G_n(a)$, although for our purposes the results obtained so far will be enough.

3 Microcanonical ensemble (NVE systems)

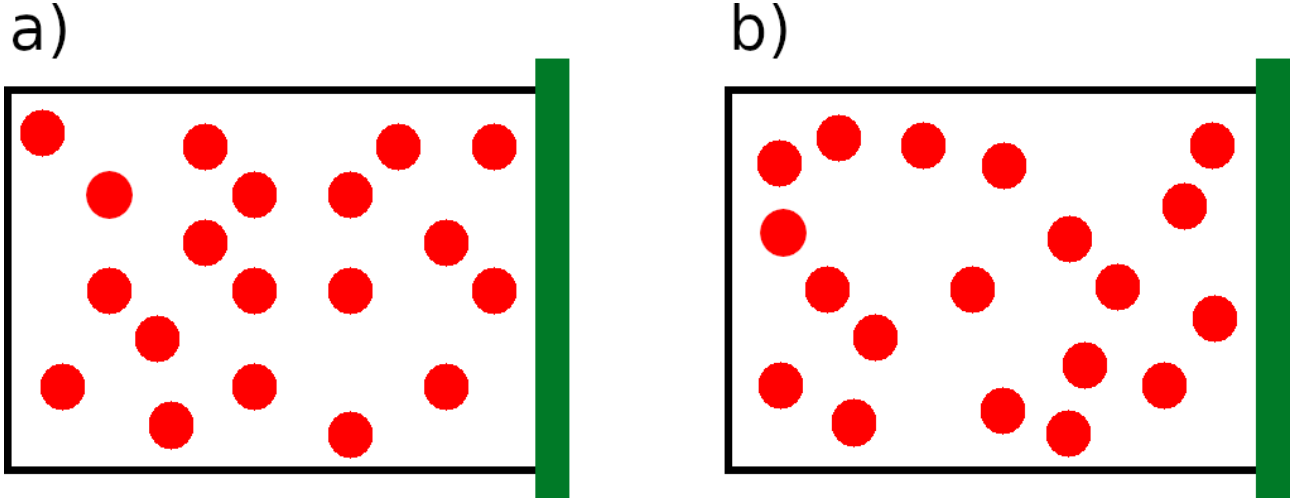


Figure 1: Two possible microscopic configurations for a box of volume V and containing N particles.

Statistical mechanics is based on the concept of ensemble. Namely, if we consider a system made of N particles under certain constraints (e.g. total volume), the *ensemble* of the system corresponds to the set of all possible microscopic configurations (e.g., positions and momenta) compatible with the macroscopic, thermodynamic state. For instance, in Fig.1 we sketch two possible microscopic configurations for an ideal gas of N particles contained in a box of volume V .

Depending on the particular choice of fixed thermodynamic variables, several possible ensembles can be considered. In the *microcanonical ensemble*, one fixes the number of particles N , the volume V and the mechanical energy E (for this reasons, one also speaks of NVE systems). As mentioned above, the idea is to derive the laws of thermodynamics based on a statistical treatment of the microstates. In order to fix the ideas, we limit our discussion to the case of a monoatomic ideal gas. The state of each particle i is characterized by its position \mathbf{r}_i and its momentum $\mathbf{p}_i = m\mathbf{v}_i$, where m is the mass of a particle (all assumed to be equal) and v_i its velocity. The microscopic state of the system is fully described by the set $\{\mathbf{r}_i, \mathbf{v}_i\}_{i=1, \dots, N}$. The constraints are implemented by imposing that \mathbf{r}_i is contained within the box of volume V and that the total energy (which for a monoatomic ideal gas is just the translational kinetic energy) equates E :

$$E = \sum_{i=1}^N \frac{1}{2} m v_i^2. \quad (19)$$

Within the microcanonical ensemble, a central assumption is the *equiprobability of the microscopic states*: under thermodynamic equilibrium, all the microscopic states have the same

probability to occur. Under this assumption, the predicted value for any macroscopic observable O can be computed as the average $\langle O \rangle$ over all possible states, as we will see below.

3.1 Entropy as the states-counting variable

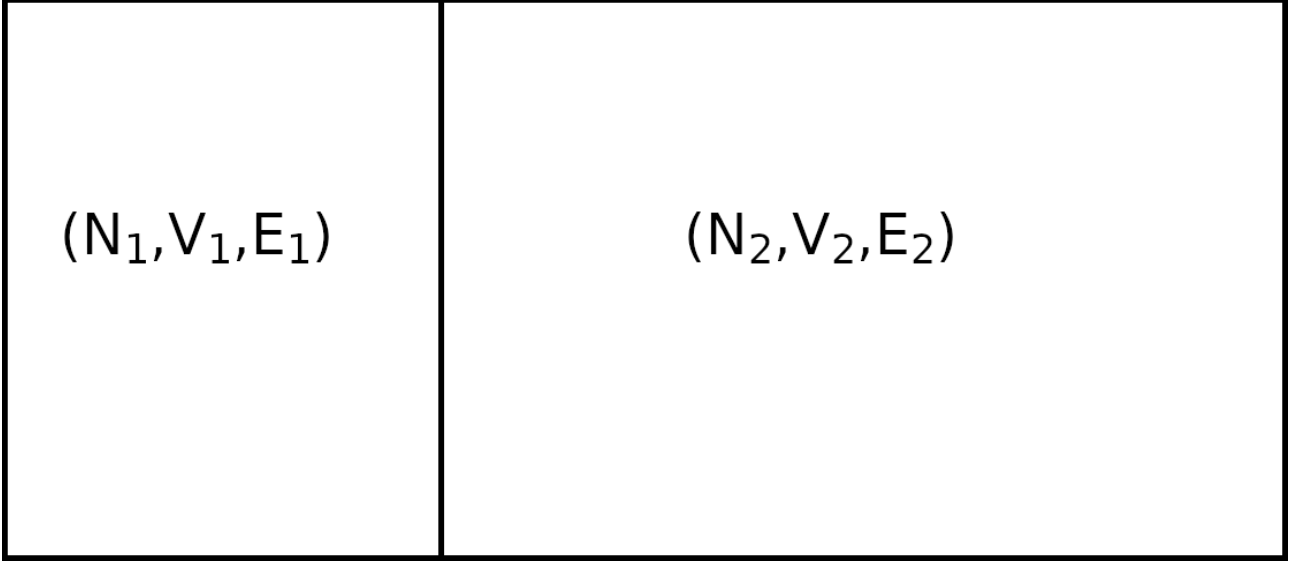


Figure 2: Two NVE systems in thermal contact.

As mentioned above, a basic assumption of the microcanonical ensemble is that all the microstates are equiprobable. To simplify the present discussion, we will assume that the system has a discrete set of microstates. If we denote by $\Omega(N, V, E)$ the number of states compatible with the NVE constraints, the probability that each of them occurs is thus $1/\Omega(N, V, E)$. Now, let us consider two microcanonical systems characterized by $(N_1, V_1, E_{1,\text{ini}})$ and $(N_2, V_2, E_{2,\text{ini}})$ and put them in thermal contact (Fig.2). The systems are allowed to exchange energy but not matter, and their volume is also kept constant. The total system formed by 1 and 2 has a total energy $E = E_1 + E_2$, which is conserved throughout the equilibration of the two systems. After some time, the two systems will have energies $E_{1,\text{fin}}$ and $E_{2,\text{fin}}$, which still satisfy $E_{1,\text{fin}} + E_{2,\text{fin}} = E$. Note that these energies are not exactly constant, so that $E_{1,\text{fin}}$ and $E_{2,\text{fin}}$ are actually the average values around which the two energies fluctuate. As we will see, the fluctuations in this system are much smaller than the average values themselves, so that this approximation turns out to be extremely precise.

Since the two systems do not exchange matter, for each microstate of system 1 at energy E'_1 (with $0 \leq E'_1 \leq E$) any microstate of system 2 at energy $E'_2 = E - E'_1$ is equally likely. The total number of states of the overall system corresponding to a given choice of E'_1 is thus

$$\Omega_{\text{glob}}(N_1, N_2, V_1, V_2, E'_1, E) = \Omega_1(N_1, V_1, E'_1) \cdot \Omega_2(N_2, V_2, E - E'_1) . \quad (20)$$

The total number of possible states \mathcal{N} in the overall system is obtained by summing over all possible values of E'_1 :

$$\mathcal{N} = \sum_{E'_1=0}^E \Omega_{\text{glob}}(N_1, N_2, V_1, V_2, E'_1, E) . \quad (21)$$

Therefore, the probability that, within the overall system, system 1 has energy E'_1 is $p_1(E'_1, E) = \Omega_{\text{glob}}(N_1, N_2, V_1, V_2, E'_1, E)/\mathcal{N} = \Omega_1(N_1, V_1, E'_1) \cdot \Omega_2(N_2, V_2, E'_2 = E - E'_1)$. The most probable energy is the one maximizing this probability, hence it satisfies the condition $\partial p_1(E'_1, E)/\partial E'_1 = 0$. This implies that

$$\left. \frac{\partial \Omega_1}{\partial E'_1} \right|_{E'_1=E_{1,\text{fin}}} \cdot \Omega_2|_{E_{2,\text{fin}}=E-E_{1,\text{fin}}} - \Omega_1|_{E'_1=E_{1,\text{fin}}} \cdot \left. \frac{\partial \Omega_2}{\partial E'_2} \right|_{E_{2,\text{fin}}=E-E_{1,\text{fin}}} = 0 , \quad (22)$$

where we used the fact that $E'_2 = E - E'_1 \Rightarrow \partial/\partial E'_2 = -\partial/\partial E'_1$. The previous formula can be rearranged as

$$\frac{1}{\Omega_1} \left. \frac{\partial \Omega_1}{\partial E'_1} \right|_{E'_1=E_{1,\text{fin}}} = \frac{1}{\Omega_2} \left. \frac{\partial \Omega_2}{\partial E'_2} \right|_{E_{2,\text{fin}}=E-E_{1,\text{fin}}} . \quad (23)$$

We further take advantage of the identity $(dy/dx)/y = d \ln y/dx$ and write

$$\left. \frac{\partial \ln \Omega_1}{\partial E'_1} \right|_{E'_1=E_{1,\text{fin}}} = \left. \frac{\partial \ln \Omega_2}{\partial E'_2} \right|_{E_{2,\text{fin}}=E-E_{1,\text{fin}}} . \quad (24)$$

Each side of the previous formula involves a quantity computed on only one of the subsystems. Therefore, this equation states that at thermal equilibrium this quantity has the same value for both systems. Hence, by the zeroth law of thermodynamics, the quantity $(\partial \ln \Omega/\partial E)_{N,V}$ has to be intimately related with temperature. In this regard, we note that one of Maxwell relations reads $(\partial E/\partial S)_{N,V} = T \Rightarrow (\partial S/\partial E)_{N,V} = 1/T$. Written in this form, this Maxwell relation closely resembles $(\partial \ln \Omega/\partial E)_{N,V}$. It is thus tempting to identify the entropy with $\ln \Omega$ up to a multiplicative, universal constant. In order to obtain an entropy from a dimensionless number, this constant has to have the units of energy divided by temperature, which is a condition satisfied by Boltzmann's constant k_B . We thus identify

$$S = k_B \ln \Omega . \quad (25)$$

We note that this definition nicely captures both the fact that S is a state variable (since the number of states are determined by the thermodynamic variables N, V, E) and that S is additive, as it should be for an extensive quantity. The latter is easily demonstrated, since for a composite system such as the one reported in Fig.2 the total number of states is $\Omega = \Omega_1 \cdot \Omega_2$, hence $S = k_B \ln \Omega = k_B(\ln \Omega_1 + \ln \Omega_2) = S_1 + S_2$. This definition of entropy is extremely important and insightful. It is important because it enables linking the macroscopic world (thermodynamics) with the microscopic features of a system. It is insightful because it gives a clear and intuitive *definition of S as the variable quantifying how many microscopic configurations are compatible with a given thermodynamic state.*

3.2 Pressure and chemical potential in NVE systems

We now consider the same setup as in the previous section, with the difference that now the wall is isolated but mobile. The same procedure as above can be followed, although now the variable quantity is the volume of each subsystem. We thus have

$$\Omega_{\text{glob}}(N_1, N_2, V'_1, V, E_1, E_2) = \Omega_1(N_1, V'_1, E_1) \cdot \Omega_2(N_2, V - V_1, E_2) . \quad (26)$$

Following the same steps as above, we end up with

$$\left. \frac{\partial \ln \Omega_1}{\partial V'_1} \right|_{V'_1=V_{1,\text{fin}}} = \left. \frac{\partial \ln \Omega_2}{\partial V'_2} \right|_{V_{2,\text{fin}}=V-V_{1,\text{fin}}} . \quad (27)$$

Since the two subsystems are at mechanical equilibrium, we expect a close relationship between $(\partial \ln \Omega / \partial V)_{N,E}$ and pressure. Since $S = k_B \ln \Omega$, we can write

$$\left(\frac{\partial \ln \Omega}{\partial V} \right)_{N,E} = \frac{1}{k_B} \left(\frac{\partial S}{\partial V} \right)_{N,E} . \quad (28)$$

Now, from the first law, in the differential form, we have $dE = TdS - pdV \Rightarrow dS = dE/T + pdV/T \Rightarrow (\partial S / \partial V)_{N,E} = p/T$. Hence

$$\left(\frac{\partial \ln \Omega}{\partial V} \right)_{N,E} = \frac{p}{k_B T} . \quad (29)$$

Analogously, we can also fix the wall to its position but enable particle exchange. We thus obtain

$$\Omega_{\text{glob}}(N'_1, N, V_1, V_2, E'_1, E) = \Omega_1(N'_1, V_1, E'_1) \cdot \Omega_2(N - N'_1, V, E - E'_1) . \quad (30)$$

Note that there is no way in which one can exchange particles without exchanging energy: if a particle goes from system 1 to system 2, it will also bring its kinetic energy with it! However, the total energy is still maintained, and we can still consider the energies with the maximum likelihood as done above. Hence, we maximize separately the probability for the energy and the one for the number of particles, obtaining in the latter case

$$\left. \frac{\partial \ln \Omega_1}{\partial N'_1} \right|_{N'_1=N_{1,\text{fin}}} = \left. \frac{\partial \ln \Omega_2}{\partial N'_2} \right|_{N_{2,\text{fin}}=N-N_{1,\text{fin}}} . \quad (31)$$

Since the two subsystems are at chemical equilibrium, we expect a close relationship between $(\partial \ln \Omega / \partial N)_{V,E}$ and the chemical potential μ . Since $S = k_B \ln \Omega$, we can write

$$\left(\frac{\partial \ln \Omega}{\partial N} \right)_{V,E} = \frac{1}{k_B} \left(\frac{\partial S}{\partial N} \right)_{V,E} . \quad (32)$$

We now consider the first law in differential form and including the particle exchange. We have $dE = TdS - pdV + \mu dN \Rightarrow dS = dE/T + pdV/T - \mu dN/T \Rightarrow (\partial S/\partial N)_{V,E} = -\mu/T$. Hence

$$\left(\frac{\partial \ln \Omega}{\partial N}\right)_{V,E} = -\frac{\mu}{k_B T}. \quad (33)$$

From these definitions, all the rest of thermodynamics follows. For instance, the Helmholtz free energy is $F = E - TS$, the Gibbs free energy is $G = F + pV$, etc.

3.3 Probabilistic concept of irreversibility

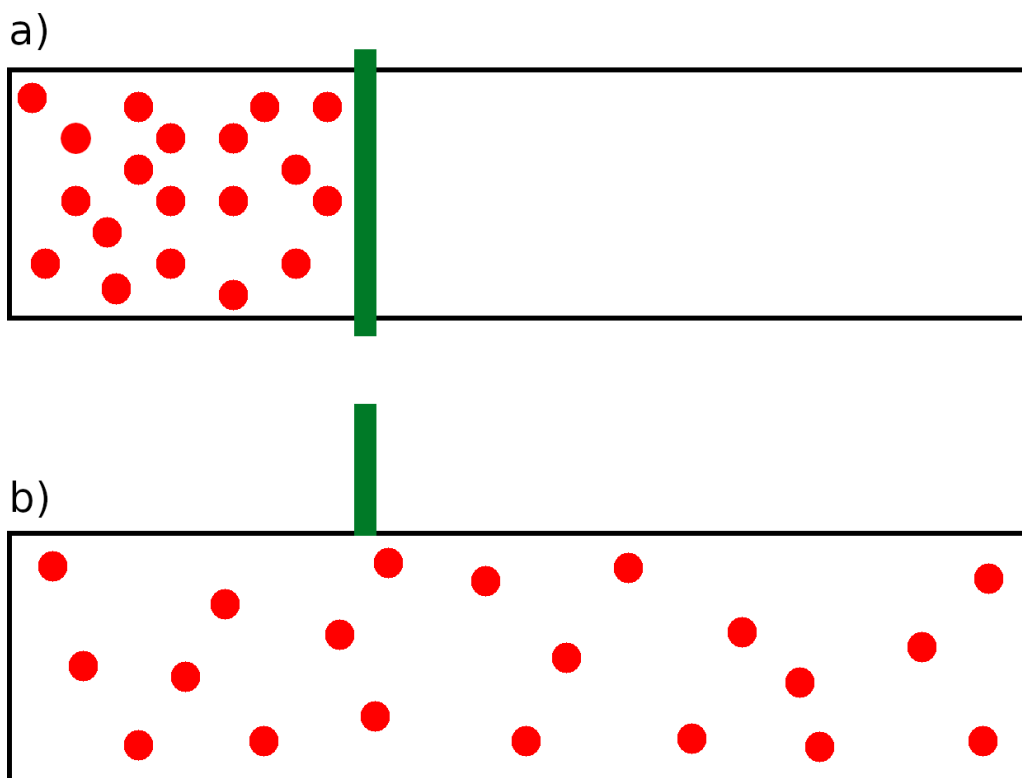


Figure 3: a) A gas of particles is initially contained in a certain volume V , with a number of microstates equal to Ω . b) By removing the wall, the gas is now free to diffuse into a volume $V' > V$, with a corresponding number of states $\Omega' > \Omega$.

The microscopic interpretation of entropy as a states-counting variable (Eq.25) provides a clear understanding of the concept of irreversibility. To this aim, in Fig.3a we consider a gas at equilibrium contained in a box of volume V . The corresponding number of states is Ω . If we remove the green wall, the gas diffuses occupying a volume $V' > V$ (Fig.3b), whose number of states is Ω' . We note that the states compatible with the system in the original volume V

are still available. However, now the particles have a lot more states compatible with the new conditions, so that $\Omega' > \Omega$. From a microscopic perspective, it is not impossible that all the particles spontaneously go back into the volume V . The probability of occurrence of this event is Ω/Ω' . Nevertheless, it turns out that there is an enormous difference in orders of magnitude between Ω' and Ω , so that this probability is practically zero. That is, if at any time we decide to close again the wall, the probability that by chance all the particles are found on the left side is $\Omega/\Omega' \simeq 0$. Therefore, the microscopic picture gives a clear and intuitive understanding of *irreversibility*, where this concept is not strictly forbidden but extremely improbable.

3.4 The Gibbs Paradox

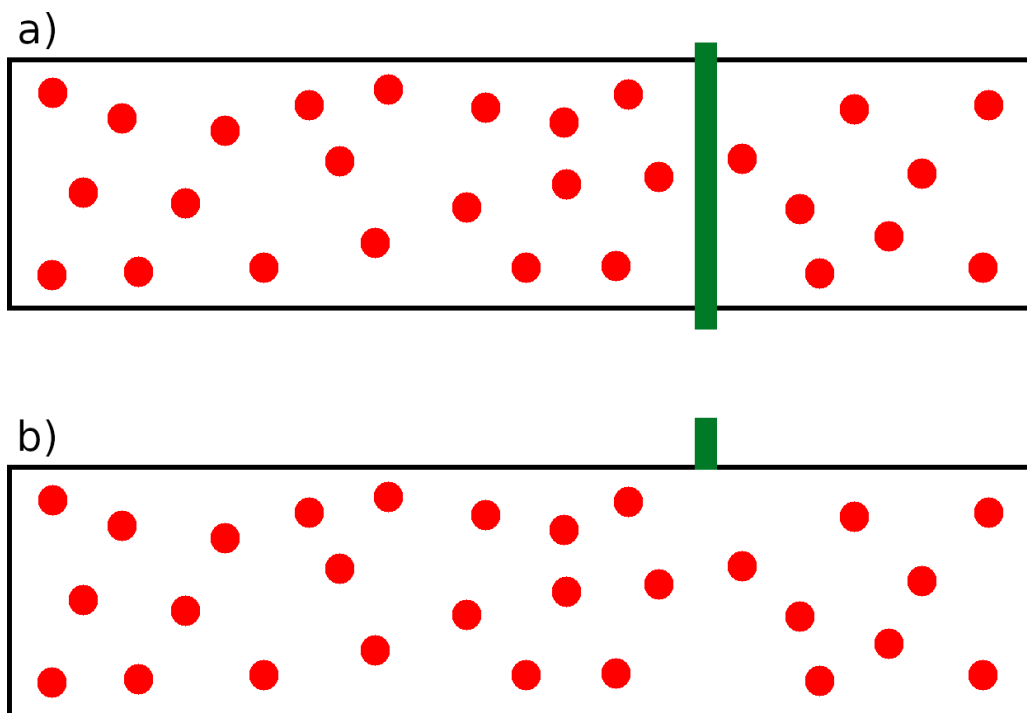


Figure 4: a) Two systems containing the same gas separated by a mobile, conducting wall are in thermal and mechanical equilibrium. b) Removing the wall does not change the macrostate of the overall system.

The Gibbs paradox is an ideal experiment that shows the importance of particle indistinguishability. If we consider the system in Fig.3b and close again the wall, we will find a microscopic state such as the one depicted in Fig.4a. For the present discussion, we will neglect the velocities for clarity, but the argument can include them without additional troubles. From the point of view of thermodynamics, the macroscopic state of the overall system is exactly the same. Now, let us assume that the particles are distinguishable and let us denote as $\omega_{1,i}$ the

number of states available to particle i in box 1. Assuming that we are dealing with an ideal gas of identical particles, the number of states does not depend on i , i.e. $\omega_{1,i} = \omega_1$. Moreover, since the particles are independent from each other, the total number of states in box 1 is $\Omega_1 = \omega_1^{N_1}$, where N_1 is the number of particles in box 1. Analogously, the total number of states in box 2 is $\Omega_2 = \omega_2^{N_2}$. Therefore, when the wall is present the total number of states is $\Omega = \Omega_1 \cdot \Omega_2 = \omega_1^{N_1} \omega_2^{N_2}$. Now, let us remove the wall (Fig.4b). As mentioned above, the thermodynamic state is still the same, since there will be no changes in total number of particles, pressure or temperature. However, the total number of states available to a single particle will be $\omega > \omega_1, \omega_2$, where the inequality holds because any microstate being counted e.g. by ω_1 is also contributing to ω . The total number of states will now be $\Omega' = \omega^{N_1+N_2} > \omega_1^{N_1} \omega_2^{N_2} = \Omega$. Since the entropy is simply $S = k_B \ln \Omega$, we thus obtain $S' > S$, which *wrongly* indicates that the two thermodynamic states reported in Fig.4 are different!

What is the problem? As mentioned above, the issue stems from considering the particles as distinguishable. Let us focus on the microstate reported in Fig.4. If the particles are distinguishable, swapping any two particles (say the ones at the top-left and bottom-left corners) results into a *different* microstate. This despite the fact that the “picture” remains exactly the same. What happens if we include indistinguishability? Under this assumption, microstates deriving from exact swapping of particles are considered to be the same. How many swappings are there? We have N particles to which we have to assign a set of N coordinates (and velocities, although in the present discussion we are neglecting them). The number of possible assignments is equal to the number of permutations of N objects, which is equal to $N!$ (see Eq.(1)). Therefore, the actual number of states in Fig.4a is $\Omega_1 = \omega_1^{N_1}/N_1!$ and $\Omega_2 = \omega_2^{N_2}/N_2!$, i.e. $\Omega = \omega_1^{N_1} \omega_2^{N_2}/(N_1! N_2!)$. When the wall is removed, the number of states are $\Omega' = \omega^{N_1+N_2}/(N_1 + N_2)!$.

In order to check whether this fixed the issue, we need to have a closer look at ω . To this aim, let us consider a system with N_0 particles in a volume V_0 . We divide the volume into small boxes of size v , hence for each particle there are $\omega_0 = V_0/v$ boxes in total (Fig.5). Accounting for the indistinguishability, the total number of states is thus

$$\Omega_0 = \frac{\omega_0^{N_0}}{N_0!} = \frac{V_0^{N_0}}{N_0! v^{N_0}} . \quad (34)$$

Indicating the number density of the gas as $\rho_0 = N_0/V_0$, we have $V_0 = N_0/\rho_0$. Moreover, according to Stirling’s approximation, $N_0! \simeq N_0^{N_0}/e^{N_0}$ for $N_0 \gg 1$, so that we obtain

$$\Omega_0 \simeq \frac{N_0^{N_0}}{N_0! \left(\frac{\rho v}{e}\right)^{N_0}} = \left(\frac{e}{\rho_0 v}\right)^{N_0} . \quad (35)$$

Armed with this result and observing that, due to mechanical equilibrium, $\rho_1 = \rho_2 = \rho$, we thus obtain for our original problem

$$\Omega = \Omega_1 \Omega_2 \simeq \left(\frac{e}{\rho_1 v}\right)^{N_1} \left(\frac{e}{\rho_2 v}\right)^{N_2} = \left(\frac{e}{\rho v}\right)^{N_1+N_2} \quad (36)$$

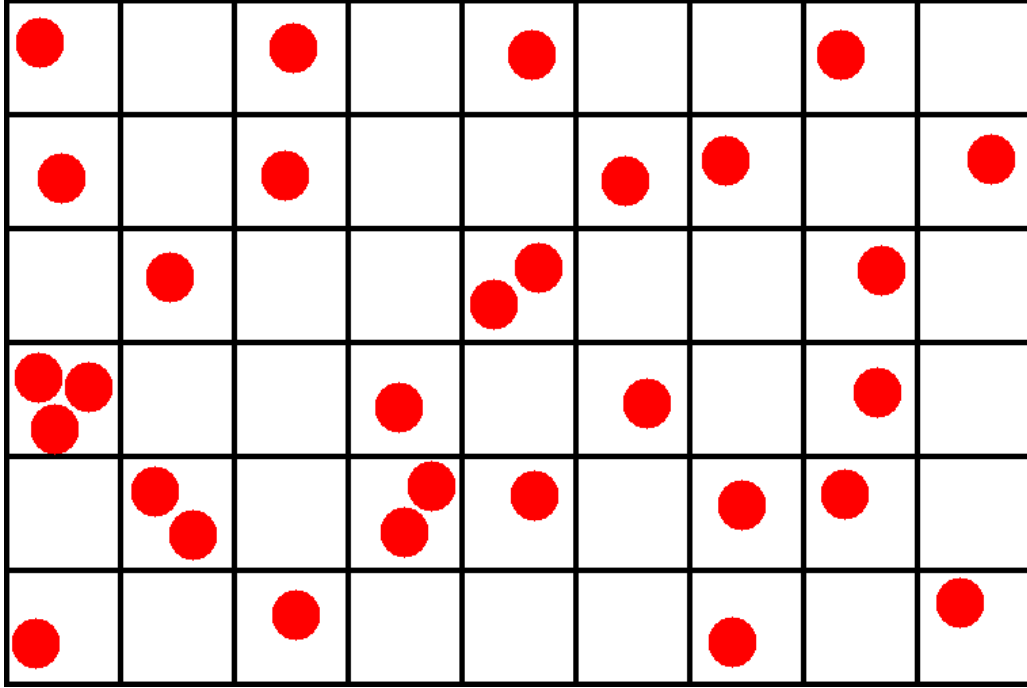


Figure 5: Example of microstate in a system with a discretized volume.

and

$$\Omega' = \left(\frac{e}{\rho v} \right)^{N_1 + N_2}. \quad (37)$$

Therefore, $\Omega' = \Omega \Rightarrow S' = S$ which solves the paradox! Therefore, *accounting for the indistinguishability of particles is paramount to correctly compute the entropy of a system.*

3.5 Entropy of mixing

Let us consider again the system in Fig.4, where now the particles in box 1 are different than the particles in box 2 (Fig.6a). When the wall is removed, the particles of both species are free to occupy the whole volume of the composite system. However, particles belonging to different species *are* distinguishable. That is, swapping the coordinates of a red and a blue particle in Fig.6 *does* result into a different microstate. As a consequence, the number of microstates in Fig.6b is larger than in Fig.6a, so that after opening the wall here will be an increase in entropy which is called *entropy of mixing*. Qualitatively, this is clearly indicated by the irreversibility of the process. Indeed, if after mixing we close again the wall, the system will not be reverted back to the state corresponding to Fig.6a, since there will be particles of both species on both sides.

Quantitatively, let N_1, V_1 and N_2, V_2 be the number of molecules within each box and the

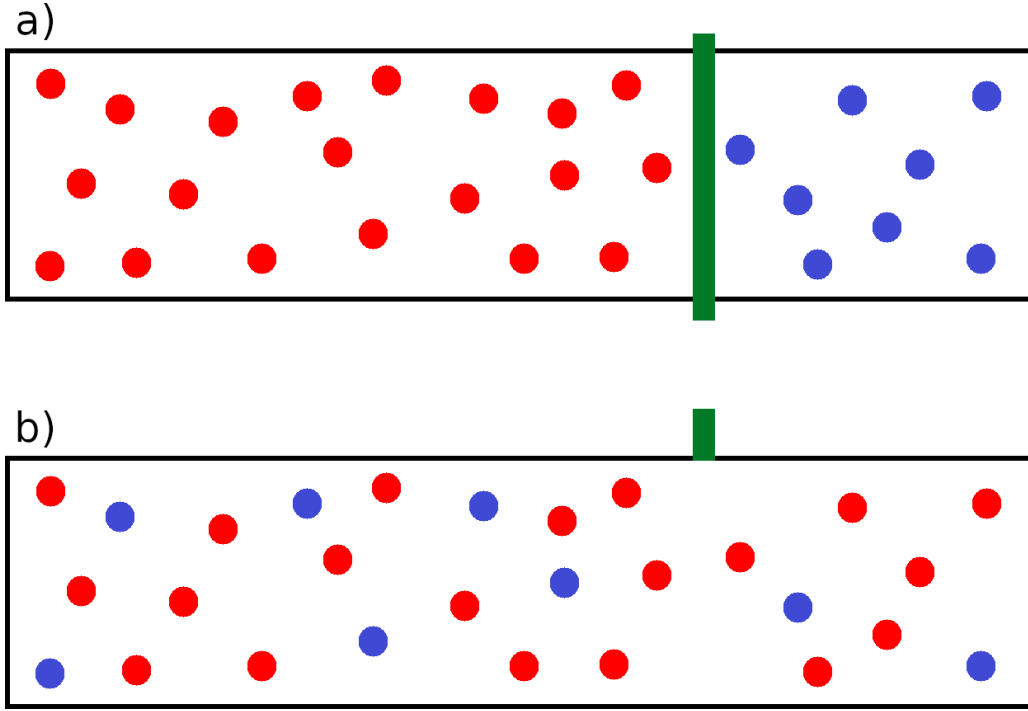


Figure 6: a) Two systems containing different species of gas separated by a mobile, conducting wall are in thermal and mechanical equilibrium. b) Since the two gases are different, removing the wall changes the macrostate of the overall system.

box volumes, respectively. The corresponding number densities before mixing are $\rho_1 = N_1/V_1$ and $\rho_2 = N_2/V_2$. Using Eq.(35), the total number of states in Fig.6a is

$$\Omega = \left(\frac{e}{\rho_1 v} \right)^{N_1} \left(\frac{e}{\rho_2 v} \right)^{N_2}. \quad (38)$$

After mixing, the density of the gases will decrease to the values $\rho'_1 = N_1/(V_1 + V_2) = \rho_1 V_1/(V_1 + V_2)$ and $\rho'_2 = N_2 V_2/(V_1 + V_2)$. The corresponding number of states is thus

$$\Omega' = \left(\frac{e}{\rho'_1 v} \right)^{N_1} \left(\frac{e}{\rho'_2 v} \right)^{N_2} = \left(\frac{e}{\rho_1 v} \right)^{N_1} \left(\frac{e}{\rho_2 v} \right)^{N_2} \left(\frac{V_1 + V_2}{V_1} \right)^{N_1} \left(\frac{V_1 + V_2}{V_2} \right)^{N_2}. \quad (39)$$

The change in entropy upon mixing is

$$\Delta S_{\text{mix}} = k_B \ln \Omega' - k_B \ln \Omega = k_B \ln \frac{\Omega'}{\Omega} = N_1 \ln \frac{V_1 + V_2}{V_1} + N_2 \ln \frac{V_1 + V_2}{V_2} > 0. \quad (40)$$

This entropy is the one preventing *spontaneous phase separation* between the two gases.

3.6 Ideal gas

The equation of state for the ideal gas can be derived by means of the microcanonical ensemble. Let us consider the ideal gas as a NVE system. For simplicity, we consider the monoatomic case. Differently than above, we consider the proper treatment of energy and volume as continuous variables. It is worth noting that referring to $\Omega(N, V, E)$ as the “number of states” is here somewhat imprecise due to the continuous character of the system. Nonetheless, we will keep using it in what follows. $\Omega(N, V, E)$ can be computed by integrating over the microscopic degrees of freedom $\{\mathbf{r}_i, \mathbf{v}_i\}$. The volume constraint is easily implemented by imposing that the coordinates \mathbf{r}_i lie within V . The energy constraint given by Eq.(19) is a bit more cumbersome and can be implemented by considering *Dirac’s delta function* $\delta(x)$. For the present purposes, this function can be thought of as a switch that, when used inside an integral, selects only the terms for which $x = 0$. The number of states can then be written as

$$\Omega(N, V, E) = \frac{C^N}{N!} \int_V d\mathbf{r}_1 \int_V d\mathbf{r}_2 \cdots \int_V d\mathbf{r}_N \int d\mathbf{v}_1 \int d\mathbf{v}_2 \cdots \int d\mathbf{v}_N \delta\left(\frac{1}{2}m \sum_{i=1}^N v_i^2 - E\right), \quad (41)$$

where C is a constant with dimensions of volume⁻¹ · velocity⁻¹ included to ensure that $\Omega(N, V, E)$ is adimensionalized, while the $1/N!$ accounts for the indistinguishability of particles, as seen above. The presence of Dirac’s delta in Eq.(41) ensures that only the velocities sets satisfying the energy constraints give a contribution to the integral, although it also makes the computation of the integral quite involved. Nevertheless, we note that the spatial contribution can be easily integrated out, giving $\int_V d\mathbf{r}_i = V$ for all the particles. By defining $g(E, N) \equiv (C^N/N!) \int d\mathbf{v}_1 \int d\mathbf{v}_2 \cdots \int d\mathbf{v}_N \delta(m \sum_{i=1}^N v_i^2/2 - E)$, we can thus write

$$\Omega(N, V, E) = g(E, N)V^N. \quad (42)$$

The entropy of the gas is obtained by means of Eq.(25) as

$$S(N, V, E) = k_B \ln \Omega(N, V, E) = S_0(N, E) + Nk_B \ln V, \quad (43)$$

where $S_0 = k_B \ln g(E, N)$ is independent of V . According to Maxwell’s relations, we also have

$$\frac{p}{T} = \left(\frac{\partial S}{\partial V}\right)_{N, E} = \frac{Nk_B}{V}. \quad (44)$$

Comparing the first and last side of the previous formula, we get $pV = Nk_B T$, which is the equation of state of the ideal gas.

3.7 A feeling for numbers

In our microscopic derivation of the various thermodynamic quantities, we have considered only the most likely values for energy, volume or number of particles. However, some fluctuations

are expected which could jeopardize the accuracy of this approach. In order to get a feeling for numbers, we consider the system of Fig.2 with an isolating, mobile wall. We saw that the maximum number of states corresponds to the volumes $V_{1,\text{fin}}$ and $V_{2,\text{fin}} = V - V_{1,\text{fin}}$. Assuming to have a monoatomic ideal gas, we also saw that the number of states can be written as $\Omega(N, V, E) = g(E, N)V^N$. Therefore, the number of states of the overall system corresponding to the most probable macrostate is $\Omega(N_1, V_{1,\text{fin}}, E_1) \cdot \Omega(N_2, V - V_{1,\text{fin}}, E_2) = g(E_1, N_1)g(E_2, N_2)V_{1,\text{fin}}^{N_1}V_{2,\text{fin}}^{N_2}$. For the sake of discussion, we assume that $N_1 = N_2 = N/2$ and that $E_1 = E_2 = E/2$. Under these symmetric conditions, it is evident that $V_{1,\text{fin}} = V_{2,\text{fin}} = V/2$. We further simplify the system assuming that it has the shape of a parallelepiped, with the separating surface having extension S and the total length equal to l , hence $V = S \cdot l$. The maximum-likelihood values for the length of each subsystem are $l_{1,\text{fin}} = l_{2,\text{fin}} = l/2$.

For this system, the total number of states is $\Omega(N_1, V_{1,\text{fin}}, E_1) \cdot \Omega(N_2, V - V_{1,\text{fin}}, E_2) = g(E_1, N_1)g(E_2, N_2)V_{1,\text{fin}}^{N_1}V_{2,\text{fin}}^{N_2} = g(E/2, N/2)^2 S^N l_1^{N/2} l_2^{N/2}$, with $l_2 = l - l_1$. Hence, the probability $p(l_1)$ that the subsystem 1 has a length l_1 is $p(l_1) = \text{constant} \cdot l_1^{N/2} (l - l_1)^{N/2}$, where the multiplicative constant is independent of l_1 . Now, as we said the largest probability is found for $l_1 = l/2$ and is equal to $p(l/2) = \text{constant} \cdot (l/2)^N$. In order to get an idea of the importance of fluctuations, we can compare this maximum value with the probability that a small displacement $dl \ll l$ is observed. To this aim, we thus compute the ratio

$$\frac{p\left(\frac{l}{2} + dl\right)}{p\left(\frac{l}{2}\right)} = \frac{\left(\frac{l}{2} + dl\right)^{\frac{N}{2}} \left(\frac{l}{2} - dl\right)^{\frac{N}{2}}}{\left(\frac{l}{2}\right)^N} = \frac{\left[\left(\frac{l}{2}\right)^2 - dl^2\right]^{\frac{N}{2}}}{\left[\left(\frac{l}{2}\right)^2\right]^{\frac{N}{2}}} = \left(1 - \frac{4dl^2}{l^2}\right)^{\frac{N}{2}}. \quad (45)$$

Taking the logarithm of both sides, we can write

$$\ln \frac{p\left(\frac{l}{2} + dl\right)}{p\left(\frac{l}{2}\right)} = \frac{N}{2} \ln \left(1 - \frac{4dl^2}{l^2}\right) \simeq -\frac{2Ndl^2}{l^2}, \quad (46)$$

where we used $\ln(1 + x) \simeq x$ for $x \ll 1$. Finally, we can thus write

$$\frac{p\left(\frac{l}{2} + dl\right)}{p\left(\frac{l}{2}\right)} \simeq e^{-\frac{2Ndl^2}{l^2}}. \quad (47)$$

Choosing $l = 0.1$ m and $S = 0.01$ m² gives a total volume $V = S \cdot l = 0.001$ m³ = 1 liter, which is a fair “real-world” value. As a reference for the ideal gas, we choose dry air at room temperature, which has a density $\rho \simeq 1.225$ kg/m³ and a molar mass $m_{\text{mol}} \simeq 29$ g/mol. The total number of molecules is thus $N = \rho V N_A / m \simeq 2.5 \cdot 10^{22}$, where $N_A \simeq 6.022 \cdot 10^{23}$ mol⁻¹ is Avogadro’s constant. Plugging the numbers into Eq.(47) shows that even for a very small displacement $dl=1$ Å = 10^{-10} m (which is the order of magnitude of the size of an atom), the relative probability of observing it is $p(l/2+dl)/p(l) \simeq e^{-50000}$. To have an idea of how huge this number is, we observe that the number of digits of e^{50000} is roughly $\log_{10} e^{50000} = 50000 \log_{10} e \simeq 21714$. Therefore,

the relative probability is roughly $p(l/2 + dl)/p(l) \simeq 1/10^{21714}$, that is the denominator is “1” followed by 21714 zeros!

This gives an idea of how precise is the maximum-likelihood approach for everyday systems. As mentioned in the introduction, this is however only true for macroscopic systems far from phase transitions. If the system is small or close to a phase transition, fluctuations may become important.

Problem. From the previous derivation, estimate the relative probability of a relative displacement $dl/l = 0.1$ for a system of $N = 100$ particles. More in general, it is useful to plot $p(l/2 + dl)/p(l)$ as a function of N for several choices of the relative displacement dl/l .

4 Canonical Ensemble (NVT systems)

The microcanonical ensemble has the advantage of being intuitive and is fundamental to establish a clear link between thermodynamics and the physics at the microscopic level. Nevertheless, it has limited practical utility, since most systems of interest are thermalized rather than isolated. In this regard, a more fruitful approach is provided by considering as thermodynamic constraints the number of particles N , the volume V and the temperature T , which define the *canonical ensemble* or NVT system.

4.1 Boltzmann's weight

From a practical perspective, the temperature T is maintained by putting the system in thermal contact with a *reservoir*, which by definition is much bigger than the system itself and is thus practically unaffected by the interaction. In order to derive the fundamental laws of the canonical ensemble, we proceed in a similar way as done for the microcanonical ensemble. Let N, V, E be the number of particles, the volume and the energy of the system. In the canonical ensemble, N and V are fixed once and for all, while E can fluctuate. Moreover, let N_r, V_r, E_r be the corresponding variables for the reservoir. By definition, one has $N_r \gg N, V_r \gg V, E_r \gg E$. The universe formed by the system of interest and the reservoir is isolated, i.e. it can be regarded to as microcanonical. Particularly the total energy $E_{\text{tot}} \equiv E + E_r$ is conserved. The probability $p(E)$ that the energy of the system of interest is E is proportional to the number of microstates of the universe compatible with E , i.e. $p(E) \propto \Omega(N, V, E) \cdot \Omega_r(N_r, V_r, E_r)$. Since $E_r \gg E$, we can consider a Taylor expansion of $\ln \Omega_r(N_r, V_r, E_r)$ around E_{tot} , obtaining

$$\ln \Omega_r(N_r, V_r, E_r) \simeq \ln \Omega_r(N_r, V_r, E_{\text{tot}}) + \left. \frac{\partial \ln \Omega_r(N_r, V_r, E')}{\partial E'} \right|_{E'=E_{\text{tot}}} (E_r - E_{\text{tot}}). \quad (48)$$

As we saw in our computations on the microcanonical ensemble, $\partial \ln \Omega / \partial E = 1/(k_B T)$. This, together with the observation that $E = E_{\text{tot}} - E_r$, yields

$$\ln \Omega_r(N_r, V_r, E_r) \simeq \ln \Omega_r(N_r, V_r, E_{\text{tot}}) - \frac{E}{k_B T}. \quad (49)$$

We can thus write for the probability $p(E)$

$$p(E) = \frac{\Omega(N, V, E) e^{-\frac{E}{k_B T}}}{\mathcal{Z}}. \quad (50)$$

Hence, *in the canonical ensemble the probability of occurrence of an energy E is proportional to the number of states weighted by $e^{-E/(k_B T)}$* . This exponential term is called *Boltzmann's weight*. We cannot stress enough the importance of Eq.(50). As we will see during all this course, it permeates virtually all the results obtained in statistical physics in general, including biophysics and soft matter. The combination $k_B T$ has the units of an energy and is usually

referred to as *thermal energy*. Physically it gives an estimation of the typical interaction energy between the particles of the system and the thermostat. It is customary to define the inverse of the thermal energy as $\beta \equiv 1/(k_B T)$, which enables writing the Boltzmann's weight in the more concise form $e^{-\beta E}$.

In order to compute the actual probability, we need to determine the normalization factor \mathcal{Z} . In the case in which the energy of the system takes discrete values, one has

$$\mathcal{Z}(N, V, T) = \sum_E \Omega(N, V, E) e^{-\frac{E}{k_B T}} \quad (51)$$

while in the continuous case $p(E)$ is a probability density and

$$\mathcal{Z}(N, V, T) = \int \Omega(N, V, E) e^{-\frac{E}{k_B T}} dE. \quad (52)$$

$\mathcal{Z}(N, V, T)$ is called the *partition function* of the system and has a very important physical meaning.

4.2 The Partition Function

In the simple case in which all the microstates of the system have the same energy (which we can set to $E = 0$), the partition function becomes $\mathcal{Z} = \sum \Omega$, i.e. it corresponds to the total number of states. Therefore, in this case there is a connection between \mathcal{Z} and the entropy S , namely $S = k_B \ln \mathcal{Z}$. In the general case, we note that the average energy of the system is

$$\langle E \rangle = \sum_E E \cdot p(E) = \frac{\sum_E E \cdot \Omega e^{-\beta E}}{\mathcal{Z}}, \quad (53)$$

where we dropped the dependence on N, V, T to lighten the notation. Note that in all the subsequent notes the brackets $\langle \dots \rangle$ indicate ensemble average of the quantity written inside. From the definition of \mathcal{Z} , we note that $\sum_E E \Omega e^{-\beta E} = -\partial \mathcal{Z} / \partial \beta$, hence the following relation holds

$$\langle E \rangle = -\frac{1}{\mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial \beta} = -\frac{\partial \ln \mathcal{Z}}{\partial \beta}. \quad (54)$$

$\langle E \rangle$ corresponds to the internal energy in the thermodynamic sense. Therefore, we can write $\langle E \rangle = F + TS$, where F is the Helmholtz free energy. From the Maxwell relations, we further know that $S = -(\partial F / \partial T)_{N, V}$, hence

$$-\frac{\partial \ln \mathcal{Z}}{\partial \beta} = \langle E \rangle = F + TS = F - T \left(\frac{\partial F}{\partial T} \right)_{N, V}. \quad (55)$$

Since $\beta = 1/(k_B T)$, we note that $(\partial F / \partial T)_{N, V} = (\partial F / \partial \beta)_{N, V} d\beta / dT = -1/(k_B T^2) (\partial F / \partial \beta)_{N, V}$. Therefore, the previous formula becomes

$$-\frac{\partial \ln \mathcal{Z}}{\partial \beta} = F + T \left(\frac{\partial F}{\partial \beta} \right)_{N, V} \frac{1}{k_B T^2} = F + \beta \left(\frac{\partial F}{\partial \beta} \right)_{N, V} = \left[\frac{\partial (\beta F)}{\partial \beta} \right]_{N, V}. \quad (56)$$

Up to an additive constant, the previous equation establishes a clear link between the Helmholtz free energy and the partition function, that is $-\ln \mathcal{Z} = \beta F$. Rearranging we finally obtain

$$F(N, V, T) = -k_B T \ln \mathcal{Z}(N, V, T) . \quad (57)$$

This formula stresses the fact that \mathcal{Z} is not simply a normalization constant, but is intimately connected to the thermodynamic features of the system. Knowledge of \mathcal{Z} gives access to F , from which all the thermodynamics of the system can be derived.

Note that in Eq.(51) and Eq.(52) the partition function has been written by collecting the microstates according to their energy and considering their number $\Omega(N, V, E)$. The partition function can be directly written in terms of the microscopic degrees of freedom, thus dropping the use of Ω . If we assume that the microstates are discrete and their number is \mathcal{N} , the partition function can be written on a general ground as

$$\mathcal{Z} = \sum_{\alpha=1}^{\mathcal{N}} e^{-\beta E_{\alpha}} , \quad (58)$$

where E_{α} is the energy of microstate α . The probability of occurrence of microstate α is

$$P_{\alpha} = \frac{e^{-\beta E_{\alpha}}}{\mathcal{Z}} = \frac{e^{-\beta E_{\alpha}}}{\sum_{\alpha'=1}^{\mathcal{N}} e^{-\beta E_{\alpha'}}} . \quad (59)$$

From this formula, we obtain that $k_B T \ln P_{\alpha} = -E_{\alpha} - k_B T \ln \mathcal{Z} = -E_{\alpha} + F$, where we used Eq.(57). Taking its average, we thus obtain

$$k_B T \langle \ln P_{\alpha} \rangle = -\langle E_{\alpha} \rangle + F = -TS , \quad (60)$$

from which we find for the entropy

$$S = -k_B \langle \ln P_{\alpha} \rangle = -k_B \sum_{\alpha=1}^{\mathcal{N}} P_{\alpha} \ln P_{\alpha} . \quad (61)$$

Thus, *the entropy is uniquely determined by the probability distribution of microstates*. We note that, as one approaches $T \rightarrow 0$ K, the Boltzmann weight will favor more and more the minimum-energy state (which we assume to be unique). As a consequence, in this limit $P_{\alpha} \rightarrow 1$ for the ground state and $P_{\alpha} \rightarrow 0$ for all the others, i.e. $S \rightarrow 0$, which gives the third law of thermodynamics.

Finally, we consider the explicit definition of partition function in the context of continuous microstates. Particularly, in the case in which each microstate is defined by the set of positions and velocities, we get

$$\mathcal{Z}(N, V, T) = \frac{1}{N!} \left(\frac{m}{h} \right)^{3N} \int d\mathbf{r}_1 \int d\mathbf{r}_2 \cdots \int d\mathbf{r}_N \int d\mathbf{v}_1 \int d\mathbf{v}_2 \cdots \int d\mathbf{v}_N e^{-\frac{E(\{\mathbf{r}_i, \mathbf{v}_i\}_{i=1, \dots, N})}{k_B T}} , \quad (62)$$

where $E(\{\mathbf{r}_i, \mathbf{v}_i\})_{i=1, \dots, N}$ indicates that the energy depends in general on the whole set of degrees of freedom, while the $1/N!$ has been included to account for the indistinguishability of particles. The prefactor $(m/h)^{3N}$ has been introduced to adimensionalize \mathcal{Z} . The constant h has the units of an angular momentum (hence, mass \times speed \times position). In the present context, the choice of its value is arbitrary, since it enters in the free energy as an additive constant via Eq.(57) and it thus does not affect the thermodynamics of the system. Nevertheless, it is customary to identify h with Planck's constant in order to harmonize the present definition of \mathcal{Z} with certain considerations on quantum systems.

4.3 The Ideal Gas

As a simple application showing the power of the partition function, we derive the microscopic and macroscopic features of an ideal gas, focusing in the simplest case of monoatomic particles.

4.3.1 Partition function

The energy of a microstate is given by

$$E(\{\mathbf{r}_i, \mathbf{v}_i\})_{i=1, \dots, N} = \frac{1}{2}m \sum_{i=1}^N v_i^2 . \quad (63)$$

Substituting into Eq.(62), we obtain

$$\mathcal{Z}(N, V, T) = \frac{1}{N!} \left(\frac{m}{h}\right)^{3N} \int d\mathbf{r}_1 \int d\mathbf{r}_2 \cdots \int d\mathbf{r}_N \int d\mathbf{v}_1 \int d\mathbf{v}_2 \cdots \int d\mathbf{v}_N e^{-\frac{\frac{1}{2}m \sum_{i=1}^N v_i^2}{k_B T}} , \quad (64)$$

We note that

$$e^{-\frac{\frac{1}{2}m \sum_{i=1}^N v_i^2}{k_B T}} = e^{-\frac{\frac{1}{2}m v_1^2}{k_B T}} \cdot e^{-\frac{\frac{1}{2}m v_2^2}{k_B T}} \cdot \dots \cdot e^{-\frac{\frac{1}{2}m v_N^2}{k_B T}} = \prod_{i=1}^N e^{-\frac{\frac{1}{2}m v_i^2}{k_B T}} . \quad (65)$$

This enables rearranging Eq.(64) as

$$\mathcal{Z}(N, V, T) = \frac{1}{N!} \left[\left(\frac{m}{h}\right)^3 \int d\mathbf{r}_1 \int d\mathbf{v}_1 e^{-\frac{\frac{1}{2}m v_1^2}{k_B T}} \right] \cdot \dots \cdot \left[\left(\frac{m}{h}\right)^3 \int d\mathbf{r}_N \int d\mathbf{v}_N e^{-\frac{\frac{1}{2}m v_N^2}{k_B T}} \right] . \quad (66)$$

In the previous formula, all the terms in the square brackets are exactly the same and correspond to the partition function of one particle (since $1! = 1$):

$$\mathcal{Z}(1, V, T) = \left(\frac{m}{h}\right)^3 \int d\mathbf{r} \int d\mathbf{v} e^{-\frac{\frac{1}{2}m v^2}{k_B T}} . \quad (67)$$

Hence, the partition function of the system can be rewritten as

$$\mathcal{Z}(N, V, T) = \frac{1}{N!} [\mathcal{Z}(1, V, T)]^N . \quad (68)$$

In Eq.(67), the position can be integrated out, giving $\int d\mathbf{r} = V$. As for the integral on the velocity, considering spherical polar coordinates we get

$$\int d\mathbf{v} e^{-\frac{1}{2}mv^2/k_B T} = \int_0^\infty dv \int_0^\pi d\theta \int_0^{2\pi} d\phi v^2 \sin\theta e^{-\frac{1}{2}mv^2/k_B T}, \quad (69)$$

where the factor $v^2 \sin\theta$ stems from the Jacobian of the coordinates transformation. The angular integrals are easily computed as

$$\int_0^{2\pi} d\phi = 2\pi, \quad \int_0^\pi \sin\theta d\theta = -\cos\theta|_0^\pi = 2. \quad (70)$$

The integral on the magnitude v is a Gaussian integral (compare Eq.(15)) giving

$$\int_0^\infty v^2 e^{-\frac{1}{2}mv^2/k_B T} dv = \left(\frac{2k_B T}{m}\right)^{\frac{3}{2}} \frac{\sqrt{\pi}}{4}. \quad (71)$$

Collecting these results, we thus obtain

$$\int d\mathbf{v} e^{-\frac{1}{2}mv^2/k_B T} = \left(\frac{2\pi k_B T}{m}\right)^{\frac{3}{2}}. \quad (72)$$

Therefore, the single-particle partition function reads

$$\mathcal{Z}(1, V, T) = \left(\frac{m}{h}\right)^3 V \left(\frac{2\pi k_B T}{m}\right)^{\frac{3}{2}} = V \left(\frac{2\pi m k_B T}{h^2}\right)^{\frac{3}{2}} \quad (73)$$

The overall partition function thus becomes

$$\mathcal{Z}(N, V, T) = \frac{V^N}{N!} \left(\frac{2\pi m k_B T}{h^2}\right)^{\frac{3N}{2}}. \quad (74)$$

4.3.2 Maxwell-Boltzmann distribution

We now consider the infinitesimal probability $\bar{p}(\mathbf{r}, \mathbf{v}) d\mathbf{r} d\mathbf{v}$ that a particle has position and velocity in the neighborhood of \mathbf{r} and \mathbf{v} , respectively. By definition, one has:

$$\bar{p}(\mathbf{r}, \mathbf{v}) d\mathbf{r} d\mathbf{v} \propto \frac{e^{-\frac{E(\mathbf{r}, \mathbf{v})}{k_B T}}}{\mathcal{Z}(1, V, T)} d\mathbf{r} d\mathbf{v}, \quad (75)$$

where $E(\mathbf{r}, \mathbf{v})$ is the corresponding energy. In the previous formula, we have considered a proportionality rather than an equality because integrating the probability one has to find 1, which is not guaranteed according to the choice of the normalization constant for $\mathcal{Z}(1, V, T)$.

As a matter of fact, this is indeed what happens in the present case, as we will see shortly. In the case of an ideal gas, $E(\mathbf{r}, \mathbf{v}) = mv^2/2$. Hence, making use of Eq.(73), we have

$$\bar{p}(\mathbf{r}, \mathbf{v})d\mathbf{r}d\mathbf{v} \propto \frac{e^{-\frac{1}{2}\frac{mv^2}{k_B T}}}{V \left(\frac{2\pi mk_B T}{h^2}\right)^{\frac{3}{2}}} d\mathbf{r}d\mathbf{v} . \quad (76)$$

Since the energy does not depend on the position, the latter can be integrated out. The probability density $p(\mathbf{v})$ for the velocity is thus

$$p(\mathbf{v}) \propto \frac{e^{-\frac{1}{2}\frac{mv^2}{k_B T}}}{\left(\frac{2\pi mk_B T}{h^2}\right)^{\frac{3}{2}}} . \quad (77)$$

Note that this is the probability density of *the vector* \mathbf{v} , although it depends only on its magnitude according to a Gaussian function. Particularly, it is easy to show that on average $\langle \mathbf{v} \rangle = 0$, which is expected from the isotropy of the system. Let us denote as A the proportionality constant, i.e.

$$p(\mathbf{v}) = A \frac{e^{-\frac{1}{2}\frac{mv^2}{k_B T}}}{\left(\frac{2\pi mk_B T}{h^2}\right)^{\frac{3}{2}}} . \quad (78)$$

In order to find A , we switch to spherical polar coordinates and integrate out the angular degrees of freedom (which gives a factor 4π), thus obtaining

$$p(v) = 4\pi A v^2 \frac{e^{-\frac{1}{2}\frac{mv^2}{k_B T}}}{\left(\frac{2\pi mk_B T}{h^2}\right)^{\frac{3}{2}}} , \quad (79)$$

where the factor v^2 stems from the Jacobian of the transformation to polar coordinates. Note that, in contrast to Eq.(78), in the previous formula we are considering the probability density of the *magnitude* of the velocity. As we will see shortly, despite the fact that $\langle \mathbf{v} \rangle = 0$ as a vector, one has $\langle v \rangle > 0$ for the magnitude. This should not come as a surprise, since $v \geq 0$ by definition and the condition $\langle v \rangle = 0$ is fulfilled only if $v = 0$ for every particle. The value of A is determined by imposing that the integral of the probability density is equal to 1:

$$1 = \int_{v=0}^{\infty} p(v)dv = \int_{v=0}^{\infty} 4\pi A v^2 \frac{e^{-\frac{1}{2}\frac{mv^2}{k_B T}}}{\left(\frac{2\pi mk_B T}{h^2}\right)^{\frac{3}{2}}} dv . \quad (80)$$

Making use of Eq.(15), we obtain from the previous equation that $A = m^3/h^3$. Substitution into Eq.(78) and Eq.(79) finally gives

$$p(\mathbf{v}) = \left(\frac{m}{2\pi k_B T}\right)^{\frac{3}{2}} e^{-\frac{mv^2}{2k_B T}} \quad (81)$$

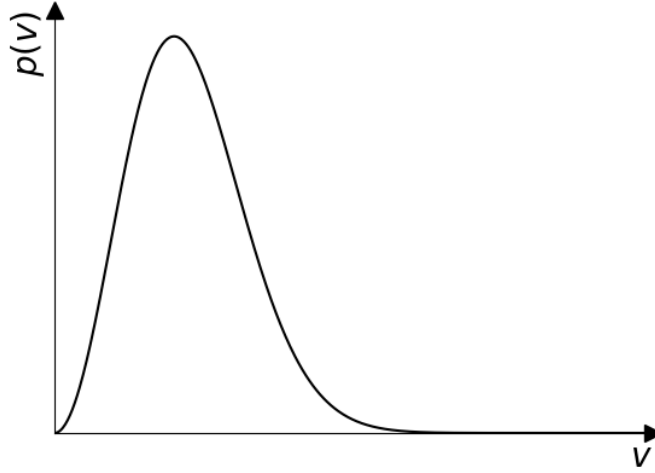


Figure 7: Maxwell-Boltzmann distribution for the magnitude of the velocities.

and

$$p(v) = 4\pi v^2 \left(\frac{m}{2\pi k_B T} \right)^{\frac{3}{2}} e^{-\frac{mv^2}{2k_B T}}, \quad (82)$$

respectively. Eq.(81) is called the *Maxwell-Boltzmann distribution* of velocities. In Fig.7, we plot the probability density for the magnitude of the velocities as obtained from Eq.(82). The plot shows a sharp peak around a certain maximum value, which thus characterizes a typical velocity of the particles. In this regard (although not exactly equivalent), let us compute the average squared velocity:

$$\langle v^2 \rangle = \int_0^\infty v^2 \cdot p(v) dv = \int_0^\infty 4\pi v^4 \left(\frac{m}{2\pi k_B T} \right)^{\frac{3}{2}} e^{-\frac{mv^2}{2k_B T}}. \quad (83)$$

Making use of Eq.(18), we obtain

$$\langle v^2 \rangle = 4\pi \left(\frac{m}{2\pi k_B T} \right)^{\frac{3}{2}} \frac{3\sqrt{\pi}}{8 \cdot \left(\frac{m}{2k_B T} \right)^{\frac{5}{2}}} = \frac{3k_B T}{m}. \quad (84)$$

Rearranging the previous formula, we thus have

$$\frac{1}{2} m \langle v^2 \rangle = \frac{3}{2} k_B T, \quad (85)$$

which is the same result we obtained from the kinetic theory of gases. Once again, we see *the microscopic interpretation of temperature, according to which it is proportional to the average*

kinetic energy per particle. Finally, we note that a good estimation for the typical velocity of a particle is $\sqrt{\langle v^2 \rangle} = \sqrt{3k_B T/m}$.

Problem. What is the value v_{\max} corresponding to the peak in Fig.7? Is it larger or smaller than $\sqrt{\langle v^2 \rangle}$? What does this say about the skewness of the distribution (i.e. the asymmetry with respect to the peak)?

4.3.3 Free energy and thermodynamics

From Eq.(57) and Eq.(74), we can compute the Helmholtz free energy of the monoatomic ideal gas:

$$F = -k_B T \ln \left[\frac{V^N}{N!} \left(\frac{2\pi m k_B T}{h^2} \right)^{\frac{3N}{2}} \right]. \quad (86)$$

We further consider Stirling's approximation for the factorial, $N! \simeq (N/e)^N$, so that

$$F = -k_B T \ln \left[e^N \left(\frac{V}{N} \right)^N \left(\frac{2\pi m k_B T}{h^2} \right)^{\frac{3N}{2}} \right]. \quad (87)$$

Now, we aim at separating the contributions coming from volume and temperature. Since the argument of a logarithm has to be an adimensionalized number, we introduce an arbitrary reference volume v_0 , which for instance can be assigned the typical volume occupied by a molecule, or can alternatively be ascribed to the size of the small boxes considered in Fig.5. In this way, we rewrite the previous formula as

$$F = -k_B T \ln \left[e^N \left(\frac{V}{N v_0} \right)^N \left(\frac{2\pi m k_B T v_0^{\frac{3}{2}}}{h^2} \right)^{\frac{3N}{2}} \right]. \quad (88)$$

By dimensional analysis, it is easy to show that $h^2/(2\pi m k_B v_0^{2/3}) \equiv T_0$ has the dimensions of a temperature. Now, we divide the various multiplicative terms found within the logarithm and obtain

$$F = -k_B T \ln e^N - k_B T \ln \left(\frac{V}{N v_0} \right)^N - k_B T \ln \left(\frac{T}{T_0} \right)^{\frac{3N}{2}}, \quad (89)$$

that is

$$F = -N k_B T - N k_B T \ln \frac{V}{N v_0} - \frac{3}{2} N k_B T \ln \frac{T}{T_0}. \quad (90)$$

This formula gives that F is extensive, as it should. In this regard, we note that the inclusion of $1/N!$ in the definition of the partition function is essential, since otherwise in the second term we would have had $-N k_B T \ln(V/v_0)$, which is not extensive.

From Maxwell's relations, we have that

$$S = - \left(\frac{\partial F}{\partial T} \right)_{N,V} = N k_B \ln \frac{V}{N v_0} + \frac{3}{2} N k_B \ln \frac{T}{T_0} + \frac{5}{2} N k_B, \quad (91)$$

which is known as *Sackur-Tetrode equation*. Using another Maxwell relation, we also find

$$p = - \left(\frac{\partial F}{\partial V} \right)_{N,T} = \frac{Nk_B T}{V} \Rightarrow pV = Nk_B T, \quad (92)$$

which is (once again) the equation of state of the ideal gas.

4.4 The equipartition theorem

Eq.(84) relates the average squared velocity of a particle to the thermal energy $k_B T$. Looking at its derivation, this formula is ultimately rooted in the fact that the velocity enters the energy of the system by contributing to its kinetic energy. This is a particular case of the so-called *equipartition theorem*, which is the focus of the present section.

On a general level, let us consider a system whose energy is $E(q_1, q_2, \dots, q_N)$ depends on a certain set of continuous scalar variables q_1, q_2, \dots, q_N . Let us assume that the dependence of E on q_1 is only via a quadratic term, independent of the rest of coordinates, i.e.

$$E(q_1, q_2, \dots, q_N) = \frac{1}{2} k q_1^2 + E'(q_2, q_3, \dots, q_N). \quad (93)$$

Let us further assume that the variables can take any real value. The partition function can be written as

$$\mathcal{Z} = \int_{-\infty}^{\infty} dq_1 \int_{-\infty}^{\infty} dq_2 \cdots \int_{-\infty}^{\infty} dq_N e^{-\beta E(q_1, q_2, \dots, q_N)}. \quad (94)$$

Since q_1 enters the energy only via the quadratic term, its contribution to the partition function can be separated from the rest:

$$\mathcal{Z} = \int_{-\infty}^{\infty} dq_1 e^{-\frac{\beta k q_1^2}{2}} \cdot \int_{-\infty}^{\infty} dq_2 \cdots \int_{-\infty}^{\infty} dq_N e^{-\beta E'(q_2, q_3, \dots, q_N)}. \quad (95)$$

Making use of Eq.(12) (note that the integration here goes from $-\infty$ to ∞ , so that we have to multiply by 2) and denoting $\mathcal{Z}' \equiv \int dq_2 \cdots \int dq_N e^{-\beta E'}$, we thus find

$$\mathcal{Z} = \sqrt{\frac{2\pi k_B T}{k}} \mathcal{Z}'. \quad (96)$$

Now, the average value of the energy associated to q_1 is

$$\left\langle \frac{1}{2} k q_1^2 \right\rangle = \frac{1}{2} k \frac{\int_{-\infty}^{\infty} dq_1 \int_{-\infty}^{\infty} dq_2 \cdots \int_{-\infty}^{\infty} dq_N q_1^2 e^{-\beta E(q_1, q_2, \dots, q_N)}}{\mathcal{Z}}. \quad (97)$$

Observing that also in the case of the numerator we can separate the contribution from q_1 and substituting Eq.(96), we obtain

$$\left\langle \frac{1}{2} k q_1^2 \right\rangle = \frac{1}{2} k \frac{\left(\int_{-\infty}^{\infty} dq_1 q_1^2 e^{-\frac{\beta k q_1^2}{2}} \right) \mathcal{Z}'}{\sqrt{\frac{2\pi k_B T}{k}} \mathcal{Z}'} = \frac{1}{2} k \cdot \frac{\frac{\sqrt{\pi}}{2 \left(\frac{k}{2k_B T} \right)^{\frac{3}{2}}}}{\sqrt{\frac{2\pi k_B T}{k}}} = \frac{1}{2} k_B T, \quad (98)$$

where we made use of Eq.(15) (again multiplied by 2). Isolating the first and last side of the chain of equalities, we thus get

$$\left\langle \frac{1}{2}kq_1^2 \right\rangle = \frac{1}{2}k_B T . \quad (99)$$

This is the essence of the *equipartition theorem*: *any scalar degree of freedom participating in the energy solely via a quadratic term contributes an average energy equal to $k_B T/2$.*

Problem. Does the result of Eq.(99) change if the variable q goes from 0 to ∞ ? Does it change in general if q ranges from q_a to q_b ?

Based on the equipartition theorem, we can obtain very fastly the result reported in Eq.(84). Indeed, we note that the velocity enters the energy via the kinetic term $mv^2/2$. However, v is the norm of a three-dimensional vector, so that in order to apply the equipartition theorem we need to write the kinetic energy based on the scalar coordinates v_x, v_y, v_z . The kinetic energy is thus written as $mv_x^2/2 + mv_y^2/2 + mv_z^2/2$. According to Eq.(99), the thermal average of each of these terms is $k_B T/2$, so that

$$\frac{1}{2}m \langle v_x^2 \rangle = \frac{1}{2}m \langle v_y^2 \rangle = \frac{1}{2}m \langle v_z^2 \rangle = \frac{1}{2}k_B T \Rightarrow \langle v_x^2 \rangle = \langle v_y^2 \rangle = \langle v_z^2 \rangle = \frac{k_B T}{m} . \quad (100)$$

Since $v^2 = v_x^2 + v_y^2 + v_z^2$, we thus get $\langle v^2 \rangle = 3k_B T/m$, which is the result reported in Eq.(84). Note that this result goes beyond the simple case of an ideal gas. Indeed, even in the case in which the molecules interact with each other (e.g. via volume exclusion, electrostatics, etc.), Eq.(84) holds as long as the interaction energy does not depend on the velocity, which is the case for most cases of interest. Hence, *the microscopic interpretation of temperature as being proportional to the average kinetic energy per particle holds in general, not only for the case of an ideal gas.*

As a final note, we also mention explicitly the fact that the equipartition theorem can be used to determine the average potential energy of the system, as long as its hypotheses are verified. This has vast applications, since whenever one is able to write the energy of a system as the sum of independent harmonic terms, the equipartition theorem quantitatively gives the average contribution of each of them to the energy.

4.5 Ergodicity

Up to this moment, we have considered ensembles, which are the collections of all the possible microstates of a system. Assuming that it is possible to observe the microstates of a system (as it is the case in simulations), this means in practice that one obtains thermodynamic information by considering a large sample of identical copies of the system, from which the information sought is obtained as an ensemble average. However, it is also true that one particular realization will explore different microstates in time. The *ergodic hypothesis* (which holds for many systems of interest) states that, if one waits long enough, a single copy of the system will explore the vast majority of the microstates. Therefore, according to this

statement *the ensemble average and the time average of any observable are equivalent*. This means that from a practical point of view one can just focus on a particular realization of the system and follow it in time. This is a diffused approach in computational biophysics, where simulations often follow just one realization of the system of interest (e.g. a protein). Also from the experimental point of view, recent techniques such as Atomic-Force Microscopy, Optical Tweezers, Magnetic Tweezers, Förster Resonance Energy Transfer (to cite a few) enable following biomolecules one at a time, enabling the access to many features of their microstates.

4.6 Effective free energy

We have seen that the knowledge of the partition function enables deriving all the thermodynamic properties of the system. Nevertheless, in biophysics it is pretty rare to have access to this quantity. From an analytical perspective, apart from the simplest models it is usually impossible to compute \mathcal{Z} . In Molecular Dynamics simulations, one has access to all the positions and velocities of the atoms composing the system. Hence, by running a long simulation, one could in principle compute \mathcal{Z} numerically for the given thermodynamic constraints used. However, it turns out that a precise estimation of the partition function via this “brute force” approach is impossible, since with the present computational power it involves extremely long simulation times. From an experimental point of view, it is also impossible to gain knowledge on the behavior of all the coordinates of a system even via single-molecule techniques, where one gathers only some partial information.

In order to both obtain a better insight on a system and to properly interpret experimental information, it is useful to introduce *collective variables* and their associated *effective free energy*. A collective variable is some global combination of the microscopic variables, which has often a direct experimental counterpart. One example is provided by the total energy E , which we have already implicitly considered as a collective variable when we introduced the partition function in Eq.(52). Indeed, considering a system made of N particles characterized by their positions \mathbf{r}_i and velocities \mathbf{v}_i , we said that the partition function of the system is given by Eq.(62). What is the connection between these two equations? In general the energy is a certain function $f(\{\mathbf{r}_i, \mathbf{v}_i\}_{i=1, \dots, N})$. Hence, one can think of ideally separating the microstates (defined by the whole set $\{\mathbf{r}_i, \mathbf{v}_i\}_{i=1, \dots, N}$) in different bins according to the value of their energy. That is, in the bin associated to the energy E we collect all the microstates satisfying $E = f(\{\mathbf{r}_i, \mathbf{v}_i\}_{i=1, \dots, N})$. This enables writing the partition function according to Eq.(52), where Ω is given by the number of microstates being contained in the bin corresponding to E . The only caveat to consider in this case is that we are dealing with continuous quantities, so that “number of microstates” is an imprecise wording, as it rather corresponds to a suitable integral over the coordinates. Formally, this binning operation can be performed by means of the Dirac delta. Eq.(52) can

be rearranged with the present notation as

$$\begin{aligned}\mathcal{Z} &= \frac{1}{N!} \left(\frac{m}{h}\right)^{3N} \int d\mathbf{r}_1 \int d\mathbf{r}_2 \cdots \int d\mathbf{r}_N \int d\mathbf{v}_1 \int d\mathbf{v}_2 \cdots \int d\mathbf{v}_N \left\{ \int dE e^{-\frac{E}{k_B T}} \delta(f(\{\mathbf{r}_i, \mathbf{v}_i\}_{i=1, \dots, N}) - E) \right\} \\ &= \int dE \left\{ \frac{1}{N!} \left(\frac{m}{h}\right)^{3N} \int d\mathbf{r}_1 \int d\mathbf{r}_2 \cdots \int d\mathbf{r}_N \int d\mathbf{v}_1 \int d\mathbf{v}_2 \cdots \int d\mathbf{v}_N \delta(f(\{\mathbf{r}_i, \mathbf{v}_i\}_{i=1, \dots, N}) - E) \right\} e^{-\frac{E}{k_B T}}.\end{aligned}$$

This has the same form as Eq.(52), with

$$\Omega(N, V, E) = \frac{1}{N!} \left(\frac{m}{h}\right)^{3N} \int d\mathbf{r}_1 \int d\mathbf{r}_2 \cdots \int d\mathbf{r}_N \int d\mathbf{v}_1 \int d\mathbf{v}_2 \cdots \int d\mathbf{v}_N \delta(f(\{\mathbf{r}_i, \mathbf{v}_i\}_{i=1, \dots, N}) - E).$$

By means of the previous formula, Eq.(52) is thus an example in which the partition function has been written in terms of the collective variable E . We note that the number of variables in the integrand has passed from $6N$ ($\{\mathbf{r}_i, \mathbf{v}_i\}_{i=1, \dots, N}$) to 1 (E). Another important point is that Eq.(52) can be rewritten as

$$\mathcal{Z}(N, \mathcal{V}, \mathcal{T}) = \int dE \Omega(N, V, E) e^{-\frac{E}{k_B T}} = \int e^{-\frac{E - k_B T \ln \Omega(N, V, E)}{k_B T}} dE. \quad (101)$$

The term $k_B \ln \Omega$ is formally analogous to the total entropy of a system as obtained in the micro-canonical ensemble. Hence, it can be interpreted as an *effective* entropy $\bar{S}(N, V, E)$ associated to the subspace of microstates of energy E . Following this analogy, we can thus write

$$\mathcal{Z}(N, \mathcal{V}, \mathcal{T}) = \int e^{-\frac{E - T\bar{S}(N, V, E)}{k_B T}} dE = \int e^{-\frac{\bar{F}(N, V, E)}{k_B T}} dE, \quad (102)$$

where $\bar{F}(N, V, E) = E - T\bar{S}(N, V, E)$ is the *effective free energy* associated to the subspace of microstates of energy E . It is worth noting that, since the overall free energy of the system is $F = -k_B T \ln \mathcal{Z}$, we have the interesting relation

$$e^{-\frac{F(N, V, T)}{k_B T}} = \int e^{-\frac{\bar{F}(N, V, E)}{k_B T}} dE, \quad (103)$$

that is, the Boltzmann weight of the overall free energy is the sum of the Boltzmann weights of the effective free energies associated to a given collective variable.

The energy is just one possible example of a collective variable, whose definition depends on the system under study and on what are the physical features that we want to highlight. We will see later in the course some practical examples in biophysical systems.

5 Other ensembles

5.1 Isobaric-Isothermic Ensemble

We have hitherto focused on two ensembles:

- The microcanonical ensemble, where one fixes the number of particles N , the volume V and the total energy E . This ensemble is a useful starting point to obtain some fundamental results, but it is of limited practical usefulness.
- The canonical ensemble, where one fixes N , V and the temperature T , which is by far the most used approach.

There are other possible ensembles that can be considered, depending on which variables are being fixed. An example is provided by the *isobaric-isothermic ensemble*, also called *Gibbs ensemble* or NpT system. As suggested by the name, here the fixed thermodynamic variables are the number of particles N , the pressure p and the temperature T .

From a practical point of view, the Gibbs ensemble is obtained by putting the system in contact with a reservoir of energy and volume, so that both thermal and mechanical energy can be exchanged in order to maintain T and p . Proceeding in the same way as for the canonical ensemble, we observe that the universe formed by the system and the reservoir is an NVE system. If we indicate as N, V, E the number of particles, volume and energy of the system, and as N_r, V_r, E_r the corresponding variables in the reservoir, we thus have that N, N_r are constant (no matter exchange between the system and the reservoir), and that $V + V_r = V_{\text{tot}}$ and $E + E_r = E_{\text{tot}}$ are also constant. The probability $p(E, V)$ that the energy of the system of interest is E and the volume is V is proportional to the number of microstates of the universe compatible with E , i.e. $p(E) \propto \Omega(N, V, E) \cdot \Omega_r(N_r, V_r, E_r)$. Since $E_r \gg E$ and $V_r \gg V$, we can consider a Taylor expansion of $\ln \Omega_r(N_r, V_r, E_r)$ around E_{tot} and V_{tot} , obtaining

$$\begin{aligned} \ln \Omega_r(N_r, V_r, E_r) &\simeq \ln \Omega_r(N_r, V_{\text{tot}}, E_{\text{tot}}) + \left. \frac{\partial \ln \Omega_r(N_r, V_r, E_r)}{\partial E'} \right|_{E'=E_{\text{tot}}} (E_r - E_{\text{tot}}) \\ &+ \left. \frac{\partial \ln \Omega_r(N_r, V', E_r)}{\partial V'} \right|_{V'=V_{\text{tot}}} (V_r - V_{\text{tot}}) . \end{aligned} \quad (104)$$

As we saw in our computations on the microcanonical ensemble, $\partial \ln \Omega / \partial E = 1/(k_B T)$ and $\partial \ln \Omega / \partial V = p/(k_B T)$. This, together with the observation that $E = E_{\text{tot}} - E_r$ and $V = V_{\text{tot}} - V_r$, yields

$$\ln \Omega_r(N_r, V_r, E_r) \simeq \ln \Omega_r(N_r, V_{\text{tot}}, E_{\text{tot}}) - \frac{E}{k_B T} - \frac{pV}{k_B T} . \quad (105)$$

We can thus write for the probability $p(E, V)$

$$p(E, V) = \frac{\Omega(N, V, E) e^{-\frac{E+pV}{k_B T}}}{\mathcal{Q}} , \quad (106)$$

where the Gibbs partition function \mathcal{Q} reads (assuming a discrete set of energies and volumes for simplicity)

$$\mathcal{Q}(N, p, T) = \sum_{E, V} \Omega(N, V, E) e^{-\frac{E+pV}{k_B T}}. \quad (107)$$

\mathcal{Q} has an interesting thermodynamic interpretation. Proceeding in a similar way as we did for the canonical ensemble, we compute the average value of $E + pV$

$$\langle E + pV \rangle = \sum_{E, V} (E + pV) \cdot p(E) = \frac{\sum_{E, V} (E + pV) \cdot \Omega e^{-\beta(E+pV)}}{\mathcal{Q}}, \quad (108)$$

From the definition of \mathcal{Q} , we note that $\sum_{E, V} (E + pV) \Omega e^{-\beta(E+pV)} = -\partial \mathcal{Q} / \partial \beta$, hence the following relation holds

$$\langle E + pV \rangle = -\frac{1}{\mathcal{Q}} \frac{\partial \mathcal{Q}}{\partial \beta} = -\frac{\partial \ln \mathcal{Q}}{\partial \beta}. \quad (109)$$

$\langle E \rangle$ corresponds to the internal energy in the thermodynamic sense. Therefore, we can write $\langle E + pV \rangle = \langle E \rangle + p \langle V \rangle = F + p \langle V \rangle + TS = G + TS$, where G is the Gibbs free energy. From the Maxwell relations, we further know that $S = -(\partial G / \partial T)_{N, p}$, hence

$$-\frac{\partial \ln \mathcal{Q}}{\partial \beta} = \langle E + pV \rangle = G + TS = G - T \left(\frac{\partial G}{\partial T} \right)_{N, p}. \quad (110)$$

Since $\beta = 1/(k_B T)$, we note that $(\partial G / \partial T)_{N, p} = (\partial G / \partial \beta)_{N, p} d\beta / dT = -1/(k_B T^2) (\partial G / \partial \beta)_{N, p}$. Therefore, the previous formula becomes

$$-\frac{\partial \ln \mathcal{Q}}{\partial \beta} = G + T \left(\frac{\partial G}{\partial \beta} \right)_{N, p} \frac{1}{k_B T^2} = G + \beta \left(\frac{\partial G}{\partial \beta} \right)_{N, p} = \left[\frac{\partial (\beta G)}{\partial \beta} \right]_{N, p}. \quad (111)$$

Up to an additive constant, the previous equation establishes a clear link between the Gibbs free energy and the Gibbs partition function, that is $-\ln \mathcal{Q} = \beta G$. Rearranging we finally obtain

$$G(N, p, T) = -k_B T \ln \mathcal{Q}(N, V, T). \quad (112)$$

Hence, as in the case of the canonical ensemble, knowing the partition function enables deriving all the thermodynamics of the system.

5.2 Grand-Canonical Ensemble

Finally, we mention the existence of the *grand canonical ensemble*, in which the system is allowed to exchange matter with the reservoir, whose role is now to maintain¹ the temperature T , the volume V and the chemical potential μ . Proceeding exactly as in the case of the canonical

¹One can also build another ensemble where the pressure is fixed rather than the volume.

and Gibbs ensembles, it can be shown that the probability $p(E, N)$ of occurrence of an energy E and a number of particles N is

$$p(E, N) = \frac{\Omega(N, V, E) e^{-\frac{E}{k_B T}} e^{\frac{\mu N}{k_B T}}}{\bar{\mathcal{Q}}}, \quad (113)$$

where the grand-canonical partition function $\bar{\mathcal{Q}}$ reads (assuming a discrete set of energies for simplicity)

$$\bar{\mathcal{Q}}(\mu, V, T) = \sum_{E, N} \Omega(N, V, E) e^{-\frac{E}{k_B T}} e^{\frac{\mu N}{k_B T}}. \quad (114)$$

Splitting the sums, this can be written as

$$\bar{\mathcal{Q}}(\mu, V, T) = \sum_N \mathcal{Z}(N, V, T) e^{\frac{\mu N}{k_B T}}, \quad (115)$$

where $\mathcal{Z}(N, V, T)$ is the canonical partition function in the case of N particles. This is the basic formula from which the various computations within the grand-canonical ensemble are performed. In general, the grand-canonical ensemble is typically the approach of choice whenever the system can interchange particles with a reservoir, such as in the case of liquid-vapour equilibrium.

Further reading

- R. K. Pathria and P. D. Beale, *Statistical Mechanics*, 3rd Edition, Chapter 1, Elsevier
- K. Huang, *Statistical Mechanics*, 2nd Edition, Chapter 1, John Wiley & Sons
- R. Delgado Buscalioni, notes on Statistical Mechanics (you can find them in the Moodle)