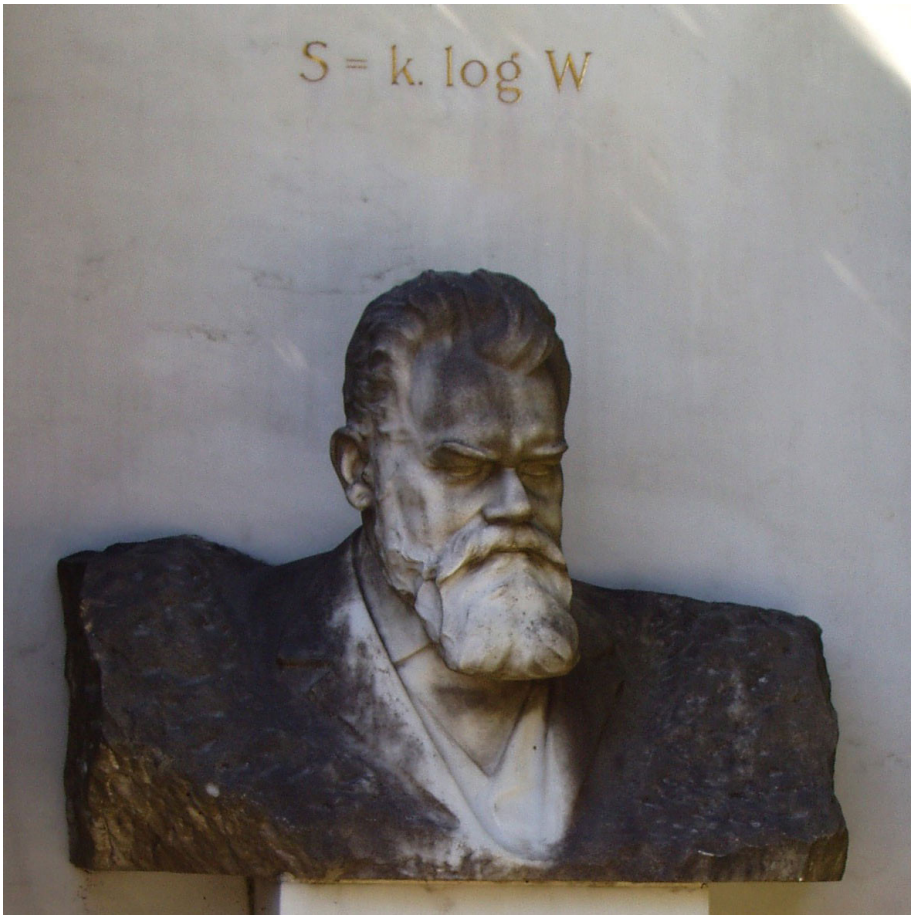
Master 1 – Space Science and Technology – AstroParis - Observatoire de Paris - PSL

# Statistical Physics

**François Levrier**

**2025 – 2026**




*The tombstone of Ludwig Boltzmann (1844-1906) in Vienna*

# Foreword

The purpose of thermodynamics is not so much to establish equations of state as to forbid some of them, for violating its core principles (the conservation of energy and the non-decreasing nature of entropy as a function of time). The purpose of statistical physics is more fundamental. It assumes that all macroscopic systems are made up of a very large number of microscopic particles obeying a set of specific laws, and aims to derive macroscopic properties (including thermodynamic principles and equations of state) from statistical considerations.

This course is meant as an introduction to the concepts, tools, and results of statistical physics. It strives to give the essentials for a master-level course, and is presented in two parts, dealing respectively with systems at equilibrium and with systems that are out of equilibrium.

In the first part, the basic ideas underlying all of this course are presented in Chapter 1. The usal statistical ensembles are presented in Chapters 2, 3, and 4. Applications to ideal quantum gas systems (either material particles or photons) are presented in Chapters 5, 6, and 7. The case of a system in equilibrium under an external field in treated in Chapter 8. Important applications of the methods of statistical physics are presented in Chapters 9 (magnetism) and 10 (heat capacity of solids).

In the second part, we begin by a discussion of the elementary laws of transport phenomena in Chapter 11, that are generalized to the linear theory of Onsager in Chapter 12. The spontaneous evolution of a system that is initially out of equilibrium is treated in the framework of classical thermodynamics in Chapter 13 and is rephrased in terms of statistical physics in Chapter 14.

In the third and final part, we discuss additional topics that are not central to the course but are of interest if the reader wishes to pursue more advanced material. In Chapter 15, we present kinetic theory, allowing for a computation of the transport coefficients and to show how hydrodynamics may be derived from statistical physics. Chapter 16 presents an introduction to the topic of stochastic processes, with an emphasis on the paradigm of random walks. Finally, Chapter **??** discusses the thermodynamics and statistical physics of gravitational systems (such as stellar clusters or galaxies) whose properties present pathological behaviours due to the long-range and unscreened nature of the interactions.

# Bibliography


— "Eléments de Physique Statistique" (B. Diu, C. Guthmann, D. Lederer, and B. Roulet)
— "Statistical Physics" (L. Landau and E. Lifchitz.)
— "Fundamentals of statistical and thermal physics" (F. Reif)
— "Statistical Mechanics" (S. K. Ma)
— "Statistical Mechanics" (R. Kubo)
— "States of Matter" (D. Goodstein)

# Contents

# Part I

# Equilibrium statistical physics

# 1

# Fundamental concepts and tools

In this chapter, we introduce the fundamental concepts of statistical physics, the notions of **microstates** and **macrostates**, **statistical ensembles** and statistical **entropy**. We illustrate these with two classic examples : the perfect paramagnetic cristal and free particles in a box.

## 1.1 From microscopic physics to macroscopic physics

Up until the middle of the XIX$^{\text{th}}$ century, all systems were thought of as **macroscopic**, and the various fields of physics (mechanics, optics, electricity,...) were considered as relatively separate topics, each obeying its own set of **physical laws** (Snell-Descartes, Newton, ...). The **atomistic** idea supported by the observations of **Brownian motion** led to the notion that there is a more fundamental layer of physics that should be tackled at the **microscopic** level, and that known physical laws observed at the macroscopic level should somehow be consequences of these more fundamental processes occuring at the microscopic scale, when the number $N$ of microscopic entities involved is large, i.e., $N \gg 1$. The reader should already have heard of **Avogadro's number**

$$N_A = 6.02214076 \times 10^{23} \tag{1.1}$$

that, by definition, is the number of particles contained in **1 mole** of the system. It is important to realize, however, that a system need not contain $N_A$ particles to be considered macroscopic. It is so if its number of particles is $N \gg 1$, so that even $10^{-8}\,\text{mol}$ is a macroscopic quantity, as it contains some $6 \times 10^{15}$ particles [1].

It is the goal of **statistical physics** to understand laws of macroscopic physics from a microscopic description. One example of such an approach would be the interpretation of the blue colour of a solution of $\text{Cu}^{2+}$ ions, but there is not always as straightforward an interpretation as this, and most often macroscopic physical laws only emerge as a **statistical** result over the large number of particles. For instance, the notion of **temperature** itself requires discussing the **probability distribution of particles velocities**, as we will see. The way statistical physics undertakes this task is not by brute force. Such an approach would mean writing the **equations of motion** for all $N$ particles in the system, specifying initial conditions for each particle, and solve this coupled system of $N \gg 1$ differential equations. This is simply impossible, and, as we will see, is not necessary. Statistical physics instead makes use

1. Systems of units also need to be adapted to whether a system is macroscopic or microscopic. For instance, lengths will be measured in Å or nm at the microscopic scale (the typical size of an atom is $\sim$1 Å), while other units ($\mu$m, mm, cm, m, km, AU, pc) will be used at the macroscopic scale. For energies, macroscopic units will be multiples of **Joules**, i.e., J, $\mu$J, kJ, while the typical microscopic energy will be in **electron-Volts** (with $1\,\text{eV} = 1.6 \times 10^{-19}\,\text{J}$).

of **probabilistic methods**, precisely based on the fact that $N \gg 1$, by computing **averages** over the $N$ particles[2] and taking advantage of the fact that for an **extensive** quantity $X$, that scales as the number of particles $N$, fluctuations scale as $\sqrt{N}$, and therefore relative fluctuations scale as

$$\frac{\Delta X}{X} \propto \frac{\sqrt{N}}{N} = \frac{1}{\sqrt{N}} \ll 1 \tag{1.2}$$

and are therefore very small. For instance, considering a small volume within a closed system, containing on average $\langle N \rangle = 10^{16}$ particles, the number of particles at any given time within this small volume varies typically by a factor $10^{-8}$, which is completely negligible. For completeness, we introduce the notion of **intensive** quantities, that do not depend on the size of system. For instance, pressure and temperature are intensive quantities.

## 1.2 Probabilities and statistical entropy

The impossibility to know the exact state[3] of a macroscopic system, i.e. of knowing for instance the exact position and velocity of each particle in a gas, requires one to approach the description of that system using the language of probabilities and statistics, i.e., of speaking of "the probability to find particle $i$ in the vicinity of position $\vec{r}_i$ to within $\mathrm{d}\vec{r}_i$". Thus, a quick reminder of the basics of this vast domain of mathematics is necessary[4].

### 1.2.1 Probabilities

#### Introduction

Consider performing $\mathcal{N}$ "identical" experiments. By "identical", we mean that the **control parameters** are the same (for instance, the energy, pressure and temperature of the gas, ...). Among these experiments, $\mathcal{N}_i$ of these will yield a result $y_i$ when measuring a given physical quantity $Y$. We define the probability $p_i$ of this result by

$$p_i = \lim_{\mathcal{N}\to\infty} \frac{\mathcal{N}_i}{\mathcal{N}} \tag{1.3}$$

If two results are mutually excluding, we have the probability to find either of these as the sum of the respective probabilities, which we write mathematically as

$$p_{i\cup j} = p_i + p_j \tag{1.4}$$

By considering the ensemble of possible results, we must have the normalization condition

$$\sum_i p_i = 1 \tag{1.5}$$

For two events that are statistically independent, for instance when measuring independent physical quantities $Y$ and $Z$, and finding respectively $y_i$ and $z_{i'}$, we have

$$p_{i\cap i'} = p_i p_{i'} \tag{1.6}$$

2. Or over the possible **configurations** of these particles
3. We will call that a **microstate** later on in this chapter.
4. The interested reader may delve into the subtleties of the domain using classical textbooks such as "Probability and Statistics" by A. Papoulis.

### Average and standard deviation

Consider a **discrete random variable** $X$, with possible outcomes $x_i$ with probabilities $p_i$. For instance, this might be the result of a dice throw. We define the **statistical average** (also called **mathematical expectation**) as

$$\langle X \rangle = \sum_i p_i x_i \tag{1.7}$$

The **standard deviation** $\Delta X$ is a measure of the typical width of the distribution of measurement results around the average. It is sometimes written $\sigma_X$. Its square $\Delta X^2$ is called the **variance**, and it is defined through

$$\Delta X^2 = \langle (X - \langle X \rangle)^2 \rangle \tag{1.8}$$

It is shown easily that this may be written as

$$\Delta X^2 = \langle X^2 \rangle - \langle X \rangle^2 \tag{1.9}$$

If we consider a function $Y = f(X)$ of this random variable, we have the following expressions for the average and standard deviation of $Y$,

$$\langle Y \rangle = \sum_i p_i f(x_i) \qquad \langle \Delta Y \rangle = \sum_i p_i f^2(x_i) - \left[ \sum_i p_i f(x_i) \right]^2 \tag{1.10}$$

Consider now a **continuous random variable**, which means that the possible outcomes of a measurement of $X$ are not discrete but continuous, $x \in \mathcal{D}$. This might be the measurement of a voltage in a circuit, for instance. We consider the probability $\mathrm{d}p$ that the outcome lies within the range $[x; x + \mathrm{d}x]$. This probability is proportionnal to $\mathrm{d}x$ and this defines the **probability density** $w(x)$ such that

$$\mathrm{d}p = w(x)\mathrm{d}x \tag{1.11}$$

In this case, the normalization condition now reads

$$\int_{\mathcal{D}} w(x)\mathrm{d}x = 1 \tag{1.12}$$

The definitions of the average and the variance are in this case

$$\langle X \rangle = \int_{\mathcal{D}} x w(x)\mathrm{d}x \qquad \Delta X^2 = \int_{\mathcal{D}} x^2 w(x)\mathrm{d}x - \left[ \int_{\mathcal{D}} x w(x)\mathrm{d}x \right]^2 \tag{1.13}$$

and for a quantity $Y$ derived from $X$ through a function, $Y = f(X)$, we have for instance the average

$$\langle Y \rangle = \int_{\mathcal{D}} f(x) w(x)\mathrm{d}x \tag{1.14}$$

## 1.2.2 Statistical entropy

### Definition

If the exact state of a system is unknown, this may be interpreted as a lack of information. The concept of **statistical entropy** is a quantifier of this ignorance. To define it, let us consider a set of

events $(e_i)$ with $i = 1, 2, \ldots, n$ and probabilities $p_i$. Obviously, we have $0 \leqslant p_i \leqslant 1$ and we have the normalization condition (1.5). The statistical entropy is defined as

$$S = -k \sum_{i=1}^{n} p_i \ln p_i \tag{1.15}$$

where $k$ is a strictly positive constant, which might be arbitrary, but for physical purposes we choose it to be equal to **Boltzmann's constant**

$$k_B = 1.38064852 \times 10^{-23}\,\mathrm{J\,K^{-1}} \tag{1.16}$$

We note that the impossible events ($p_i = 0$) do not contribute to the entropy, since $p_i \ln p_i = 0$.

### Positivity

The entropy is **positive**, since $p_i \geqslant 0$ and $\ln p_i \leqslant 0$ for all $i$.

$$S \geqslant 0 \tag{1.17}$$

It reaches its minimum null value if and only if the system's state is perfectly known, i.e., there is one $i$ for which $p_i = 1$ and $p_j = 0$ for all other events. Indeed, $\ln p_i = 0$ and $p_j \ln p_j = 0$ for $j \neq i$. The fact that $S = 0$ in that case clearly states that the state of the system is perfectly known.

### Maximum entropy

For a given number of events $n$, the **maximum value** of the entropy is reached when all events have the same probability, $p_i = 1/n$. In that case,

$$S = S_{\max} = -k_B \sum_{i=1}^{n} \frac{1}{n} \ln\left(\frac{1}{n}\right) = k_B \ln n \tag{1.18}$$

This may be shown through the method of **Lagrange multipliers**. This method searches for the extrema of the function

$$F\left\{(p_i\}\right) = S\left\{(p_i\}\right) - \lambda \left[\left(\sum_{i=1}^{n} p_i\right) - 1\right] \tag{1.19}$$

as if the variables $p_i$ *were* independent. The constraint is inserted as the factor multiplying the Lagrange multiplier $\lambda$. The extrema are found by writing

$$\frac{\partial F}{\partial p_i} = \frac{\partial S}{\partial p_i} - \lambda = -k_B \left(\ln p_i + 1\right) - \lambda = 0 \tag{1.20}$$

This leads to

$$\ln p_i = -\left(1 + \frac{\lambda}{k_B}\right) \tag{1.21}$$

which shows that all proababilites $p_i$ are identical. The fact that $p_i = 1/n$ stems then from the normalization.

**Additivity**

The entropy is also **additive**, that is, if two systems $\mathcal{S}$ and $\mathcal{S}'$ are independent and their entropies are respectively $S$ and $S'$, then the union $\mathcal{S}'' = \mathcal{S} \cup \mathcal{S}'$ of the two has an entropy

$$S'' = S + S' \tag{1.22}$$

To show this, consider two sets of events $(e_i)_{1 \leqslant i \leqslant n}$ and $\left(e'_j\right)_{1 \leqslant j \leqslant q}$ that are independent, and whose probabilities are $p_i$ and $p'_j$, respectively. The combined event $e''_{i,j} = e_i \cap e'_j$ has a probability $p''_{i,j} = p_i p'_j$ due to the statistical independence of the two sets. Therefore, the entropy of the combined system is

$$S'' = -k_B \sum_{i=1}^{n} \sum_{j=1}^{q} p''_{i,j} \ln p''_{i,j} = -k_B \sum_{i=1}^{n} \sum_{j=1}^{q} p_i p'_j \ln p_i p'_j \tag{1.23}$$

The logarithm may be written as a sum, so that

$$S'' = -k_B \left[ \sum_{i=1}^{n} \sum_{j=1}^{q} p_i p'_j \ln p_i + \sum_{i=1}^{n} \sum_{j=1}^{q} p_i p'_j \ln p'_j \right] \tag{1.24}$$

In the double sums, we may separate contributions from the two systems as

$$S'' = -k_B \left[ \sum_{i=1}^{n} p_i \ln p_i \sum_{j=1}^{q} p'_j + \sum_{j=1}^{q} p'_j \ln p'_j \sum_{i=1}^{n} p_i \right] \tag{1.25}$$

and use the normalization conditions to conclude that

$$S'' = -k_B \sum_{i=1}^{n} p_i \ln p_i - k_B \sum_{j=1}^{q} p'_j \ln p'_j = S + S' \tag{1.26}$$

## 1.3 Microstates and macrostates

A **macrostate** will be defined as the state of a system determined by its properties at a macroscopic level, e.g., pressure, temperature, total energy... The notion of **microstate** stems from the fact that a given macrostate can correspond to a great many different microstates, that would be determined by the properties of the microscopic constituents of the system (particle positions and momenta in a classical approach, or quantum states). By exchanging the momenta of two particles, for instance, the macroscopic properties of the system are obviously not changed, but these correspond to different microstates [5]. In the following, when we use the word "states", we mean microstates.

### 1.3.1 Quantum description of states

Any physical system may exist in a given number of stationary quantum states. In the quantum approach, these states are countable and can be labeled with an index $\ell$. We may write these states $|\ell\rangle$ using the **ket** notation introduced by Paul Dirac, and write $E_\ell$ for their respective energies. The fact that a given macrostate can correspond to many different microstates and that the exact microstate $|\ell\rangle$ cannot be known (due to Heisenberg's uncertainty principle, if nothing else) suggests, as we already mentioned, to use a probabilistic approach, and to consider that the various microstates $|\ell\rangle$ are

5. We assume here that the particles are **distinguishable**.

characterized by probabilities $p_\ell$. The macrostate of a system can be then thought of as a probability distribution $\{p_\ell\}$ over its possible microstates $|\ell\rangle$.

Consider a system in a given macrostate, and the probabilities $p_\ell$ of the possible microstates corresponding to it (also called **accessible** microstates). We assume that these probabilities are known, and we consider the measurement of a certain physical quantity $A$. Say the value of $A$ when the system is in microstate $\ell$ is[6] $a_\ell$. Then the expected value of $A$ in this macrostate is given by

$$\langle A\rangle = \sum_\ell a_\ell p_\ell \tag{1.27}$$

### 1.3.2 Density of states

The probabilities $p_\ell$ and measurements $a_\ell$ may often be viewed not as functions of the specific state $|\ell\rangle$ , but as functions of the energy $E_\ell$ of the state, so that we may write

$$p_\ell a_\ell = f\left(E_\ell\right) \tag{1.28}$$

and therefore the expected value of $A$ in this macrostate is

$$\langle A\rangle = \sum_\ell f\left(E_\ell\right) \tag{1.29}$$

Now, several states may have the same energy. For instance, in an ideal gas, exchanging the momenta of any two particles leaves the energy unchanged. So we may introduce the **degeneracy** $g\left(E_\ell\right)$ of the energy level $E_\ell$ by defining it as the number of different states having that energy. Thus the above equation may be rewritten as a sum over the energy levels rather than over the states

$$\langle A\rangle = \sum_{E_\ell} g\left(E_\ell\right) f\left(E_\ell\right) \tag{1.30}$$

It should be noted now that in a macroscopic system, the energy levels are very tightly packed. For instance, considering a free particle of mass $m$ in a $1\,\mathrm{cm}^3$ box, the difference $\delta E_\ell$ between two consecutive energy levels scales as $\sqrt{T/m}$ and is of the order of $1\,\mu\mathrm{eV}$ for an electron at temperatures of a few thousand K. This is much smaller than the accuracy with which the energy of the system may be measured, or the energy scale $\delta_0 E$ over which the function $f$ may significantly vary. This latter quantity may be estimated by

$$\delta_0 E = \left(\frac{1}{f}\frac{\partial f}{\partial E}\right)^{-1} \tag{1.31}$$

As a result, it is possible, from a macroscopic point of view, to consider that states are not discrete, but form a **continuum**, and we may write $\delta n(E)$ the number of states whose energies lie between $E$ and $E+\delta E$, with a "resolution" in energy such that $\delta E_\ell \ll \delta E \ll \delta_0 E$. This number of states, in the limit $\delta E \ll \delta_0 E$, scales linearly with the interval $\delta E$. We define the **density of states** $\rho(E)$ such that

$$\rho(E) = \frac{\delta n(E)}{\delta E} \tag{1.32}$$

The density of states thus gives the number of states per energy interval. Using this approach, the expected value of the measurement of $A$ may be written as

$$\langle A\rangle = \sum_\ell f\left(E_\ell\right) \simeq \int f(E)\rho(E)\mathrm{d}E \tag{1.33}$$

6. In quantum mechanical Dirac notation, we write $a_\ell = \langle\ell|A|\ell\rangle$.

A related concept is that of the **cumulative density of states** $\Phi(E)$ which gives the number of states whose energy lies below $E$. Obviously, we have

$$\delta n(E) = \Phi(E + \delta E) - \Phi(E) \tag{1.34}$$

and therefore the density of states is the derivative of $\Phi$

$$\rho(E) = \frac{\mathrm{d}\Phi}{\mathrm{d}E} \tag{1.35}$$

## 1.4 Ergodic principle

The necessity of speaking in statistical terms (averages, fluctuations,...) for macroscopic physical systems requires a discussion of what in meant by such an averaging procedure, a topic leading to introduce the **ergodic principle**.

### 1.4.1 Thermal motions and time evolution

Consider a gas of $N$ particles in a macroscopically-stationary state, within a box of volume $V$. Now, the number $n$ of particles contained in a small volume $v$ within the box varies with time, as particles enter and leave it, around the average value $\overline{n} = (v/V)N$.

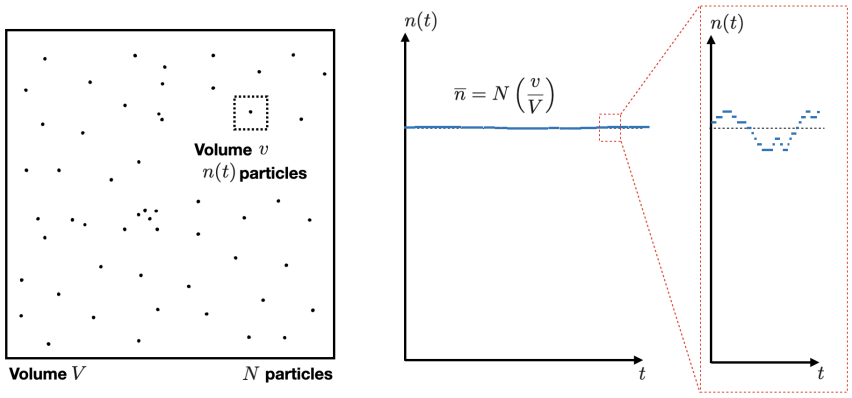



FIGURE 1.1 – Number of particles in a small but macroscopic volume, in a stationary state, and its time evolution (at large and small timescales).

This average value is computed as a time-average of $n(t)$,

$$\overline{n} = \lim_{\Delta t \to \infty} \frac{1}{\Delta t} \int_t^{t+\Delta t} n(u)\mathrm{d}u \tag{1.36}$$

In practice, the time interval $\Delta t$ need not be taken to infinity, but must be large compared to typical timescales $\delta t$ at the microscopic level (i.e., it must be so that a large number of particles have had time to leave and enter the volume $v$). As already mentioned, if $v$ is a macroscopic volume, the fluctuations of $n$ are such that $\sigma_n/\overline{n} \ll 1$.

Now consider that the gas is initially contained in half the total box, and at $t = 0$ is released to occupy all the volume. The small volume $v$ is assumed to lie in the initially empty part of the box. As time

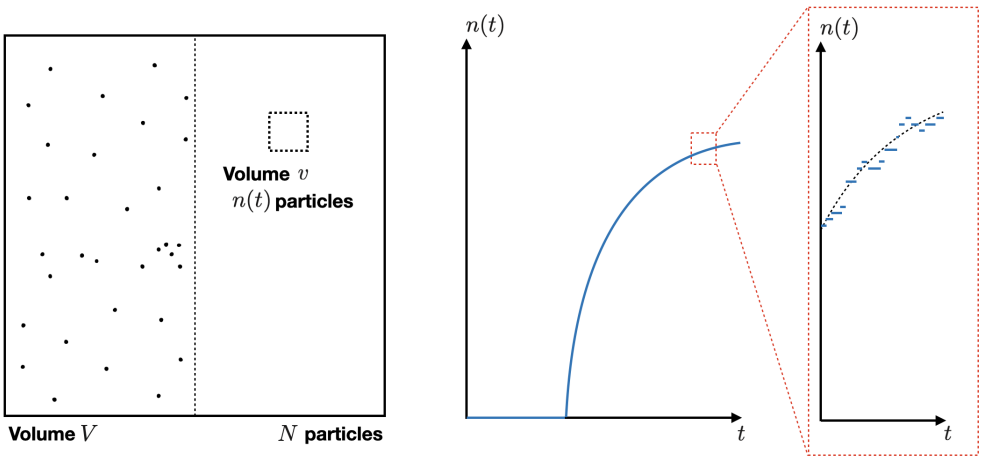



FIGURE 1.2 – Number of particles in a small but macroscopic volume, out of equilibrium, and its time evolution (at large and small timescales).

flows, particles start to enter the volume $v$, and so the curve for $n(t)$ is an increasing function of time that appears smooth at a sufficiently large timescale, but is in fact a "stair-like" function when viewed with sufficient resolution (i.e., $\delta t$ small enough). This means that there is a macroscopic timescale $\tau$ for the evolution of $n$ and the averaging timescale $\Delta t$ must be chosen such that $\delta t \ll \Delta t \ll \tau$.

### 1.4.2 Statistical ensembles

Consider the realization of a very large number $\mathcal{N}$ of identical experiments with the same macroscopic conditions and parameters, and the measurement of a physical quantity at a given point in time, say for instance the number $n$ of particles in that small volume $v$ mentioned above. This suggests that another averaging is possible, i.e.,

$$\langle n(t) \rangle = \frac{1}{\mathcal{N}} \sum_{j=1}^{\mathcal{N}} n_j(t) \tag{1.37}$$

where $n_j(t)$ is the result of the measurement of the number of particles in that small volume, at time $t$, in the experiment $j$. If the experiments have been prepared independently and without any bias, the various systems will each be in a certain state $|\ell\rangle$, and the probability of finding a given system in that state is $p_\ell(t)$. We note that this may depend on time as the example above shows. The result of the measurement of $n$ in state $|\ell\rangle$ is written $n_\ell$ and therefore

$$\langle n(t) \rangle = \sum_{\ell} p_\ell(t) n_\ell \tag{1.38}$$

The ensemble of such identical experiments is called a **statistical ensemble** and the corresponding average $\langle n \rangle$ is called **ensemble average**. We will encounter this notion in several subsequent chapters.

### 1.4.3 Ergodic principle

In many cases[7], we will use the following approach : first we determine the possible states of the system, then we determine their probabilities, which may depend on time, $p_\ell(t)$. Finally, we compute expected values of physical quantities as ensemble averages in the sense of equation (1.38). The **ergodic**

7. Such are the examples given in 1.6.

**principle** states that these ensemble averages are identical to the time averages over a single system such as written in equation (1.36), i.e.,

$$\langle A \rangle = \overline{A} \tag{1.39}$$

for any physical, macroscopic quantity $A$. This is an essential principle for statistical physics, as the theory rests on ensemble averages but experiments can in many cases only rest on time averages.

## 1.5 Classical statistical physics

Although we will mostly use a quantum approach to statistical physics in the rest of this course, it is important to realize that the original formulation of the topic was purely classical, as quantum physics was unknown to Boltzmann, of course. This classical approach will be used later to discuss kinetic theory.

### 1.5.1 Phase space and distribution function

In this classical approach of statistical physics, a microstate of a system is defined by the positions $\vec{r}_i$ and momenta $\vec{p}_i$ of the $N$ particles that the system is made of. It is therefore given by a point in a $6N$-dimensional **phase space**. The probability $\mathrm{d}P$ of finding the system in a state $\{\vec{r}_i, \vec{p}_i\}_{1 \leqslant i \leqslant N}$, or rather of finding it in an elementary volume $\mathrm{d}\tau$ around that state is then written using a **density of states**, also called **statistical distribution function**, $f_N$ as

$$\mathrm{d}P = f_N(\{\vec{r}_i, \vec{p}_i\})\mathrm{d}\tau \tag{1.40}$$

where the volume element (Fig. 1.3) is simply

$$\mathrm{d}\tau = \prod_{i=1}^{N} \mathrm{d}^3\vec{r}_i \mathrm{d}^3\vec{p}_i \tag{1.41}$$

Note that this description is valid when all **actions** (products of energy and time or of momentum and length) in the problem are much larger than the **reduced Planck constant** $\hbar$. In other words, the classical description should be recovered from the quantum one when $\hbar \to 0$.

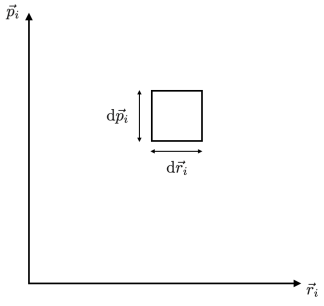



FIGURE 1.3 – An element of phase space

### 1.5.2 From the classical to the quantum

The passage from this classical description to a pseudo-quantum one usually involves dividing by $h^3$ per particle, to take into account the fact that the **Heisenberg uncertainty principle** implies that a single state of a single particle occupies a volume $\sim h^3$ in phase space. This means replacing $\mathrm{d}\tau$ by $\mathrm{d}\tau/h^{3N}$ when converting from a classical expression.

## 1.6 First examples

In this section, we apply some of the notions introduced above to give a first approach to classical problems of statistical physics, to which we will return frequently in the following chapters.

### 1.6.1 The perfect paramagnetic crystal

#### Description of the system

Consider a system of $N$ atoms, at fixed points (the nodes of the crystal's lattice), bathed in a uniform magnetic field $\vec{B} = B\vec{e}_z$. Each atom has a **spin** $s$, which we assume to be $1/2$, and associated with those spins are magnetic moments $\vec{\mu}$. Projected on the $z$ axis, this means that each atom may be in one of two states $+\mu$ or $-\mu$. Since the interaction energy is $-\vec{\mu}\cdot\vec{B}$, the "up" state $+\mu$ has an energy $\epsilon_+ = -\mu B$ and the down state $-\mu$ has an energy $\epsilon_- = \mu B$.

#### Accessible microstates

A microstate of the crystal is determined by specifying, for each atom, whether it is in the "up" or "down" state. Consequently, there are $2^N$ states, but not all of them have the same total energy. Without any interaction between the different atoms, which we assume here, the total energy is a function of one parameter, say the number $n_+$ of atoms in the "up" state,

$$E(n_+) = n_+\epsilon_+ + n_-\epsilon_- = \mu B(n_- - n_+) = \mu B(N - 2n_+) \tag{1.42}$$

using the fact that $n_+ + n_- = N$. Assuming that the total energy is fixed (the **microcanonical case**, as we will see in the next chapter), the number of accessible states, i.e., those that are compatible with the constraint of this total energy, is given by the number of possible choices of the $n_+$ "up" state atoms from a sample of $N$, so they are given by the well-known combinatorial expression

$$C_N^{n_+} = \frac{N!}{n_+!\,(N-n_+)!} \tag{1.43}$$

In the words used in the previous sections, the microstate $|\ell\rangle = |++---\ldots+-+\rangle$ is an ordered [8] succession of "up" or "down" states, the energy of the state is $E_\ell = \mu B(N - 2n_\ell)$ where $n_\ell = n_+$, and the degeneracy is $g(E_\ell) = C_N^{n_\ell}$.

#### Probabilities

Now the question is to find the probabilities $p_\ell$ for these states. Say the probability for a given atom to be in the "up" state is $\alpha$, then the probability for the $\ell$ state is [9]

$$p_\ell = \alpha^{n_\ell}(1-\alpha)^{N-n_\ell} \tag{1.44}$$

and the probability for having exactly $n$ "up" spins is

$$p(n) = C_N^n \alpha^n (1-\alpha)^{N-n} \tag{1.45}$$

8. It is ordered because the atoms are at fixed positions in space and are thus distinguishable.
9. Without any interaction between atoms, the probability $\alpha$ is the same for all atoms.

**Most probable number of up states**

We then wish to ask what the most probable value of $n$ is, and we suspect this will depend on $\alpha$. Considering that the system is macroscopic, $n$ is very large, so we may view $n$ as a continuous variable, in which case finding the most probable value $n_m$ amounts to solving

$$\frac{\partial p}{\partial n} = 0 \tag{1.46}$$

which, given that the logarithm is a monotonously increasing function, amounts to solving

$$\frac{\partial \ln p}{\partial n} = 0 \tag{1.47}$$

For large integers $m \gg 1$, we may use the **Stirling approximation** for the factorials,

$$\ln(m!) \approx m \ln m - m \qquad\qquad m! \approx \sqrt{2\pi m}\left(\frac{m}{e}\right)^m \tag{1.48}$$

This leads then to $n_m = \alpha N$, which is an unsurprising result. The derivation of it is left as an exercise.

**Fluctuations**

What is more important is to look at the approximate probability near the maximum $n_m$. By definition, we have, using the Taylor expansion formula,

$$\ln p(n) = \ln p(n_m) + (n - n_m)\left(\frac{\partial \ln p}{\partial n}\right)_{n=n_m} + \frac{(n-n_m)^2}{2}\left(\frac{\partial^2 \ln p}{\partial n^2}\right)_{n=n_m} + \dots \tag{1.49}$$

and the first derivative is zero. The second derivative is necessarily negative for $n_m$ to be a maximum. Therefore, we may introduce a quantity $\sigma_n$ through

$$\frac{1}{\sigma_n^2} = -\left(\frac{\partial^2 \ln p}{\partial n^2}\right)_{n=n_m} \tag{1.50}$$

so that near the maximum value $n_m$, the probabilities for the values of $n$ are **Gaussian-distributed**

$$p(n) \approx p\left(n_m\right) \exp\left[-\frac{(n-n_m)^2}{2\sigma_n^2}\right] \tag{1.51}$$

It is left to the reader to show that $\sigma_n = \sqrt{N\alpha(1-\alpha)}$, but the important point, already noted, is that $\sigma_n \propto \sqrt{N}$ while $n_m \propto N$, and therefore the probability to find $n$ significantly far away (i.e., a few $\sigma_n$) from the mean goes down rapidly as the size of the system increases.

$$\frac{\sigma_n}{n_m} \propto \frac{1}{\sqrt{N}} \ll 1 \tag{1.52}$$

### 1.6.2 Free particles in a box

**Description of the system**

Consider a single particle of mass $m$, free to move around in a rectangular box of size $L_x \times L_y \times L_z$. The "free" character means that there is no interaction potential applied to the particle. Its wave function $\psi\left(\vec{r}, t\right)$ obeys the **time-dependent Schrödinger equation**

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi + U\psi = -\frac{\hbar^2}{2m}\Delta\psi \tag{1.53}$$

given that $U = 0$. Obvious solutions to this equation are **plane waves** of the form [10]

$$\psi\left(\vec{r},t\right) = A\exp\left[i\left(\vec{k}\cdot\vec{r} - \omega t\right)\right] \tag{1.54}$$

provided that we have the following **dispersion relation** between the angular frequency $\omega$ and the **wavevector** $\vec{k}$, or equivalently the wavenumber $k = ||\vec{k}||$,

$$\omega = \frac{\hbar k^2}{2m} \tag{1.55}$$

The angular frequency is related to the particle's energy $\epsilon$ and the wavevector is related to its momentum through the **de Broglie relations**

$$\epsilon = \hbar\omega \qquad\qquad \vec{p} = \hbar\vec{k} \tag{1.56}$$

so that we find that the energy of the particle is just its kinetic energy, since there is no potential energy is this situation.

$$\epsilon = \frac{\hbar^2 k^2}{2m} = \frac{p^2}{2m} \tag{1.57}$$

Such plane wave solutions are stationary

$$\psi\left(\vec{r},t\right) = \phi\left(\vec{r}\right)\chi(t) \tag{1.58}$$

with the following form for the spatial and temporal parts, respectively,

$$\phi\left(\vec{r}\right) = A\exp\left(i\vec{k}\cdot\vec{r}\right) \qquad\qquad \chi(t) = \exp\left(-\frac{i\epsilon t}{\hbar}\right) \tag{1.59}$$

### Quantization

The possible values of $\vec{k}$ are those that comply with the imposed boundary conditions. With the chosen set of basis functions (complex exponentials), we must work with **periodic boundary conditions** [11], stating for instance that $\phi(0,y,z) = \phi(L_x,y,z)$. We therefore have

$$\exp\left[i\left(k_x L_x + k_y L_y + k_z L_z\right)\right] = \exp\left[i\left(k_x \times 0 + k_y L_y + k_z L_z\right)\right] \tag{1.60}$$

and so $\exp\left(ik_x L_x\right) = 1$, which means that $k_x L_x$ must be an integer multiple of $2\pi$. Applied to all three directions, this means that the wavevector must take the form

$$\vec{k} = \frac{2\pi n_x}{L_x}\vec{e}_x + \frac{2\pi n_y}{L_y}\vec{e}_y + \frac{2\pi n_z}{L_z}\vec{e}_z \tag{1.61}$$

10. These are basis solutions, the general solution is any linear combination of these, owing to the linearity of the Schrödinger equation.

11. If we had used so-called **strict boundary condtions**, where the wave function vanishes at the boundaries, we would have found sine functions with a quantization dependent on three *positive* integers $n_x$, $n_y$, and $n_z$. These strict conditions are however quite unrealistic, as is discussed in Diu *et al.*, appendix II.

where $n_x$, $n_y$, and $n_z$ are three signed integers in $\mathbb{Z}$. Obviously, in this case, a state $|\ell\rangle$ is defined through the specification of the triplet $(n_x, n_y, n_z)$, and the energy is

$$\epsilon\left(n_x, n_y, n_z\right) = \frac{\hbar^2}{2m} \times 4\pi^2 \left[ \left(\frac{n_x}{L_x}\right)^2 + \left(\frac{n_y}{L_y}\right)^2 + \left(\frac{n_z}{L_z}\right)^2 \right] \tag{1.62}$$

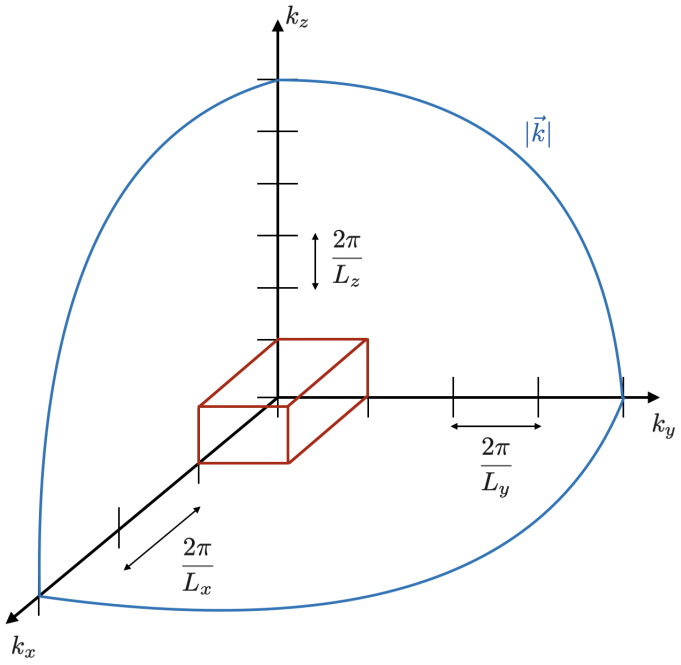



FIGURE 1.4 – Computation of the density of states for a free particle in a box

The density of states is most easily computed by considering first the cumulative density $\Phi(\epsilon)$, that is the number of states whose energy is at most $\epsilon$. Now the energy is directly related to the length of the wavevector, whose endpoints are at the nodes of a grid, and so this number is simply the ratio

$$\Phi(\epsilon) = \frac{\frac{4}{3}\pi k^3}{\frac{2\pi}{L_x} \times \frac{2\pi}{L_y} \times \frac{2\pi}{L_z}} = \frac{Vk^3}{6\pi^2} = \frac{V}{6\pi^2}\left(\frac{2m\epsilon}{\hbar^2}\right)^{3/2} = \frac{\sqrt{2}Vm^{3/2}}{3\pi^2\hbar^3}\epsilon^{3/2} \tag{1.63}$$

By derivation, we then have the density of states

$$\rho(\epsilon) = \frac{Vm^{3/2}}{\sqrt{2}\pi^2\hbar^3}\sqrt{\epsilon} \tag{1.64}$$

We note that if the particle has a spin $s$, this density of states must be multiplied by a factor $2s+1$ in order to take into account the spin states, in which case we have

$$\Phi(\epsilon) = (2s+1)\frac{\sqrt{2}Vm^{3/2}}{3\pi^2\hbar^3}\epsilon^{3/2} \qquad \rho(\epsilon) = \frac{(2s+1)Vm^{3/2}}{\sqrt{2}\pi^2\hbar^3}\sqrt{\epsilon} \tag{1.65}$$

**Extension to the case of $N$ free particles**

Now, consider $N$ such free identical particles in the same box, and assume that these do not interact, which is essentially saying that we are dealing with an ideal gas. We suppose however that these particles are **distinguishable**, i.e., that one can tell a given particle from another. This is not faithful to reality as we will see later, but leads to simpler computations. A given state is then specified by the $N$ wavevectors of the different particles, i.e.,

$$|\ell\rangle = \left|\vec{k}_1, \vec{k}_2, \ldots, \vec{k}_N\right\rangle \tag{1.66}$$

and for each wavevector we have the quantification condition

$$\vec{k}_i = \frac{2\pi n_{i,x}}{L_x}\vec{e}_x + \frac{2\pi n_{i,y}}{L_y}\vec{e}_y + \frac{2\pi n_{i,x}}{L_z}\vec{e}_z \tag{1.67}$$

so the state may also be written as $|\ell\rangle = |n_{1,x}, n_{1,y}, n_{1,z}, \ldots, n_{N,x}, n_{N,y}, n_{N,z}\rangle$, and the energy is

$$E_\ell = \frac{\hbar^2}{2m}\sum_{i=1}^{N} k_i^2 = \frac{\hbar^2 K^2}{2m} \tag{1.68}$$

where $\vec{K}$ is a $3N$-dimensional vector. The same approach used in the single-particle case can be extended, and the cumulative density of states reads

$$\Phi_N(E) = (2s+1)^N \frac{V_{3N}(K)}{\Delta v_{3N}} \tag{1.69}$$

where $V_{3N}(K)$ is the volume of the $3N$-dimensional sphere of radius $K$

$$V_{3N}(K) = C_N K^{3N} = C_N \left(\frac{2mE}{\hbar^2}\right)^{3N/2} \tag{1.70}$$

and $\Delta v_{3N}$ is the volume of the $3N$-dimensional "brick" around a given possible node for $\vec{K}$

$$\Delta v_{3N} = \frac{(2\pi)^{3N}}{V^N} \tag{1.71}$$

We have also introduced the counting of spin states for the $N$ particles by throwing in the factor $(2s+1)^N$. We therefore have

$$\Phi_N(E) = (2s+1)^N \frac{C_N V^N}{(2\pi)^{3N}} \left(\frac{2mE}{\hbar^2}\right)^{3N/2} = B_N V^N E^{3N/2} \tag{1.72}$$

and the density of states reads

$$\rho_N(E) = A_N V^N E^{3N/2-1} \tag{1.73}$$

which grows very fast as the energy increases [12].

12. For completeness, we give the full expression for this density of states, writing out the $A_N$ constant explicitly,

$$\rho_N(E) = \frac{3N}{2}\left(\frac{m}{2\pi\hbar^2}\right)^{3N/2} \frac{(2s+1)^N V^N}{\Gamma\left(\frac{3N}{2}+1\right)} E^{3N/2-1} \tag{1.74}$$

where

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} \mathrm{d}t \tag{1.75}$$

is Euler's Gamma function. This function has the following two important properties

$$\Gamma(x+1) = x\Gamma(x) \qquad \forall n \in \mathbb{N} \quad \Gamma(n+1) = n! \tag{1.76}$$

# 2 Microcanonical ensemble

In this chapter, we discuss the **microcanonical distribution**, that is the probabilities of the various microstates for a macroscopic system that is completely isolated, i.e., that cannot lose or gain energy nor particles.

## 2.1 Definitions

### 2.1.1 Control parameters

Consider a macroscopic system $\mathcal{S}$. There are some parameters affecting this system that an outside operator may be able to control, for instance its volume $V$, or the pressure $P$ that is applied to it. From a statistical point of view, these are "exact" values, not probabilistic ones, although of course they are not perfectly known and are affected by experimental uncertainties. These are called **control parameters**.

### 2.1.2 Conserved quantities

For an isolated system, some of the properties of the system are conserved, although not easily or freely imposed by the operator. Such are the energy $E$, the number of particles $N$ or the mass $M$ of the system. These keep the same value throughout the experiment, so in a sense they are akin to control parameters.

### 2.1.3 External parameters and internal variables

The control parameters and conserved quantities may be collectively called **external parameters**, and they are fixed throughout the experiment. The quantities that are free to fluctuate due to thermal motions of the particles are on the other hand called **internal variables**. Such is for instance the density of particles in the vicinity of a given point in the volume. These must be treated in a statistical way, and given probabilities to take certain values.

Depending on the situation, a given quantity may be an external parameter or an internal variable. For instance, in a rigid box, the volume $V$ is an external parameter but it becomes an internal variable if the box is closed by a piston applying a given pressure (which is then the external parameter). Similarly, if the box is thermally isolated, the energy $E$ is an external parameter, but if the box is in contact with a thermostat (see next chapter) at temperature $T$, this temperature is the external parameter and the energy becomes an internal variable.

Both external parameters and internal variables may be **extensive**, meaning that they scale linearly with the number of particles in the system, or **intensive**, meaning that they do not. By imagining

two identical systems merged into a single one, one may decide whether a variable is extensive or not. In the former category, we have the volume, the energy, and entropy for instance, while temperature, pressure, and density are examples of intensive variables. As the density $n = N/V$ shows, the ratio of two extensive variable is intensive.

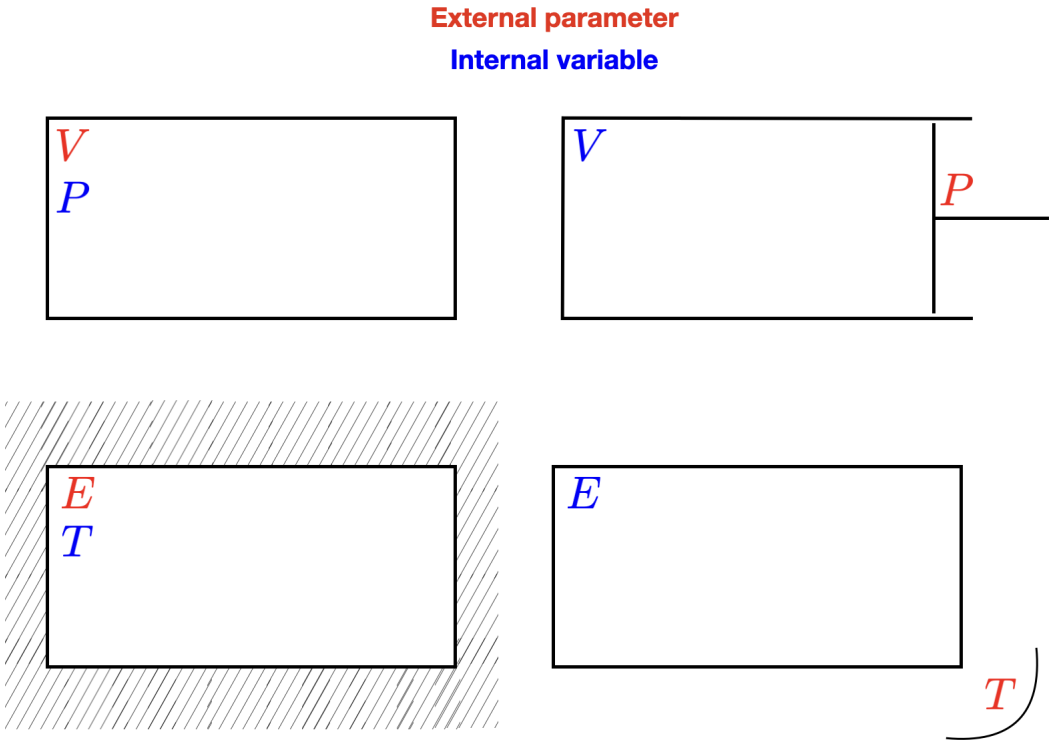



FIGURE 2.1 – External parameters and internal variables. In the top left, volume is an external parameter and pressure is an internal one. In the top right, it is the opposite. In the bottom left, energy is an external parameter and temperature is an internal one. In the bottom right, it is the opposite.

In statistical physics, it is essential to carefully define the system and the conditions that are applied to it. Depending on these choices, the proper approach may vary, as we will see in the next chapters, and the computations may be quite different.

## 2.2 The microcanonical distribution

### 2.2.1 Isolated system

A system is **isolated** if it exchanges no energy or particles with the outside world. In that case, the number $N$ of particles and the energy $E$ of the system are fixed (and in general so is the volume $V$ because the change in volume implies work done on the system, and therefore a change in its energy, generally).

The accessible microstates $|\ell\rangle$ are those for which the energy $E_\ell$ lies within the range $[E, E + \delta E]$ where $\delta E$ is the experimental uncertainty affecting the measurement of the fixed energy. The question is now to find out what the probabilities $p_\ell$ of these various microstates are.

### 2.2.2 Fundamental postulate

To answer this question, statistical physics postulates, on a fundamental level, that for an isolated system **all accessible microstates are equally probable**. Writing the number of accessible states as $\Omega(E;\delta E)$, this means that

$$E_\ell \notin [E, E+\delta E] \implies p_\ell = 0 \qquad E_\ell \in [E, E+\delta E] \implies p_\ell = \frac{1}{\Omega(E;\delta E)} \tag{2.1}$$

Given that we have no further information on the system than what its total energy is, this choice of equal probabilities seems logical (see the discussion about statistical entropy in the previous chapter). Indeed, there is no reason to think a state $|\ell\rangle$ to be privileged over another $|\ell'\rangle$ if they are both compatible with the macroscopic constraint on energy.

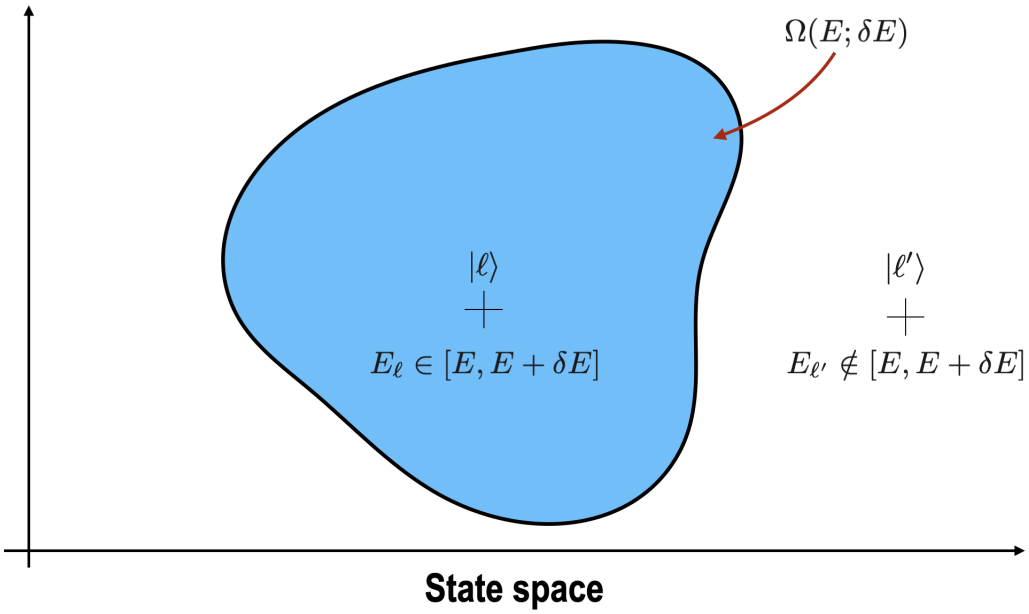



FIGURE 2.2 – The fundamental postulate states that all accessible microstates $|\ell\rangle$, in the blue-coloured subset of the space of states, are equally probable.

### 2.2.3 Microcanonical entropy

Applying the definition of the statistical entropy (1.15), we find

$$S = -k_B \sum_\ell p_\ell \ln p_\ell = -k_B \sum_\ell \frac{1}{\Omega} \ln\left(\frac{1}{\Omega}\right) \tag{2.2}$$

which leads to the famous **Boltzmann formula** giving the **microcanonical entropy**

$$S = k_B \ln \Omega \tag{2.3}$$

Now, the number of accessible states is not only a function of the energy, but also of the other external parameters, i.e., $N$, $V$, or any other parameter $x$. Introducing the density of states $\rho$ from the relation

$$\Omega(E, V, N, x; \delta E) = \rho(E, V, N, x)\delta E \tag{2.4}$$

we may also write the microcanonical entropy as

$$S = k_B \ln \rho + k_B \ln \delta E \approx k_B \ln \rho \tag{2.5}$$

Indeed, taking as an example the case of $N$ free particles in a box, we recall that $\rho_N(E) \propto E^{3N/2-1}$ and therefore the first term is linear in the number of particles and is completely dominating the second, whatever the unit used to measure energy. Of course this is only true when the density of states is a very rapidly increasing function of the energy. In fact, one can also write the entropy using the cumulative density,

$$S = k_B \ln \Phi \tag{2.6}$$

It should be stressed that the energy $E$ considered here is the **internal energy**, i.e., that of the system in a reference frame where it is at rest. In a reference frame where the system (of mass $M$) has a total momentum $\vec{P}$, the entropy is

$$S = S\left(E - \frac{\vec{P}^2}{2M}\right) \tag{2.7}$$

because the number of accessible states is independent of whether the system is macroscopically at rest or in motion.

## 2.3 Properties and behaviour of an isolated system

### 2.3.1 Statistical distribution of an internal variable

Consider an isolated system $\mathcal{S}$, with fixed energy $E$, number of particles $N$ and volume $V$. We may be interested in determining the possible values of the number $n$ of particles within a fixed volume $v$ inside the system. This is an internal variable, assigned with probabilities $p(n)$. To keep the discussion more general, we consider any internal variable $Y$, taking values $y$, and ask what are its probabilities $p(y)$. Since this internal variable may be continuous, we seek the probability density $w(y)$ such that $\mathrm{d}p(y) = w(y)\mathrm{d}y$ is the probability that $Y$ takes values between $y$ and $y + \mathrm{d}y$.

The number of accessible states for $\mathcal{S}$ is $\Omega(E)$, but among these only $\omega(E, y) = u(E, y)\mathrm{d}y$ are such that $y_\ell$ (the value of $Y$ in state $|\ell\rangle$) lies within $[y, y + \mathrm{d}y]$, as shown in Fig. 2.3. Therefore, we have

$$\mathrm{d}p(y) = \frac{\omega(E, y)}{\Omega(E)} = \frac{1}{\Omega} u(e, y)\mathrm{d}y \tag{2.8}$$

so that $w = u/\Omega$. Obviously, we have the normalisation condition

$$\int u(E, y)\mathrm{d}y = \Omega(E) \tag{2.9}$$

We may interpret $\omega(E, y)$ as the number of accessible states if, on top of $E$, $V$, $N$, ..., the variable $Y$ were also fixed at the value $y$. We may thus define a **partial microcanonical entropy**

$$s(y) = k_B \ln [\omega(E, y)] \tag{2.10}$$

which may be used to express the probability $\mathrm{d}p(y)$ as

$$\mathrm{d}p(y) \propto \omega \propto \exp\left[\frac{s(y)}{k_B}\right] \tag{2.11}$$

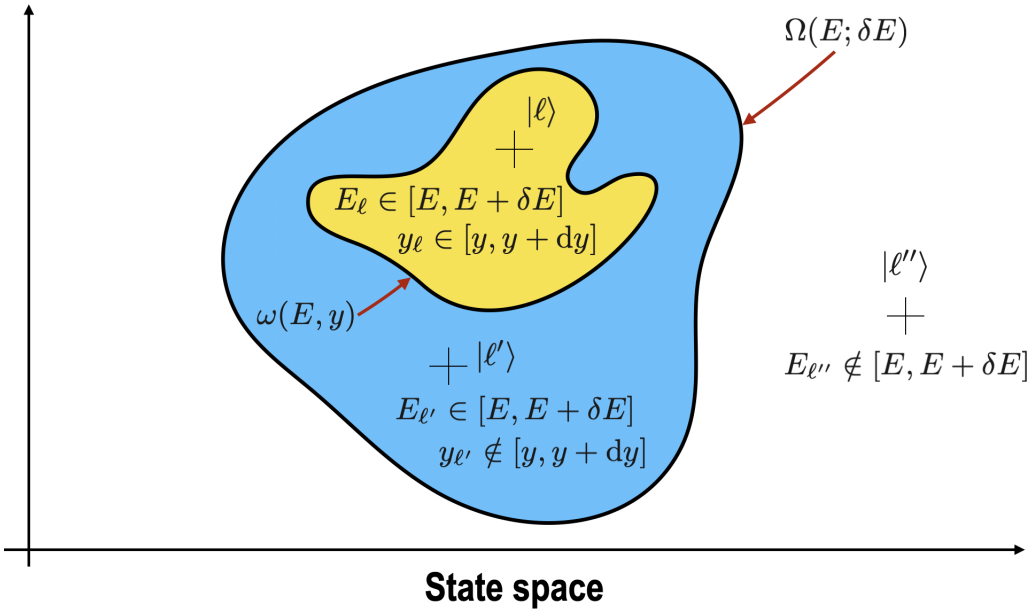



FIGURE 2.3 – While the blue ensemble is that of accessible microstates, the yellow ensemble is the subset of $\omega(E, y)$ such accessible microstates for which $y_\ell \in [y, y + \mathrm{d}y]$.

This shows that the most probable value of $Y$ is that which maximizes the partial entropy. Writing $y_m$ for this value we have the following Taylor expansion in the vicinity of $y_m$

$$s(y) = s(y_m) + \frac{(y - y_m)^2}{2} \frac{\partial^2 s}{\partial y^2}(y_m) + \dots \tag{2.12}$$

with the second derivative being negative as it is a maximum. We may therefore introduce the notation

$$\frac{1}{\sigma_y^2} = -\frac{1}{k_B} \frac{\partial^2 s}{\partial y^2}(y_m) \tag{2.13}$$

and this leads to writing the probability $\mathrm{d}p(y)$ or $w(y)$ in the vicinity of the maximum (in fact wherever the probability is not negligible) as a Gaussian distribution

$$\mathrm{d}p(y) \propto \exp\left[-\frac{(y - y_m)^2}{2\sigma_y^2}\right] \tag{2.14}$$

The value of $Y$ at the maximum is a function of the external parameters, $y_m = y_m(E, V, N, x)$. Its dependence on the number of particles may be written as $y_m \propto N^\alpha$ with $\alpha = 0$ if the variable is **intensive** and $\alpha = 1$ if the variable is **extensive**. Entropy is itself extensive, so the width of the Gaussian scales as

$$\sigma_y \propto \left(\frac{\partial^2 s}{\partial y^2}\right)^{-1/2} \propto \left(\frac{N}{N^{2\alpha}}\right)^{-1/2} \propto N^{\alpha - 1/2} \tag{2.15}$$

The relative fluctuations then scale as

$$\frac{\sigma_y}{y_m} \propto \frac{1}{\sqrt{N}} \tag{2.16}$$

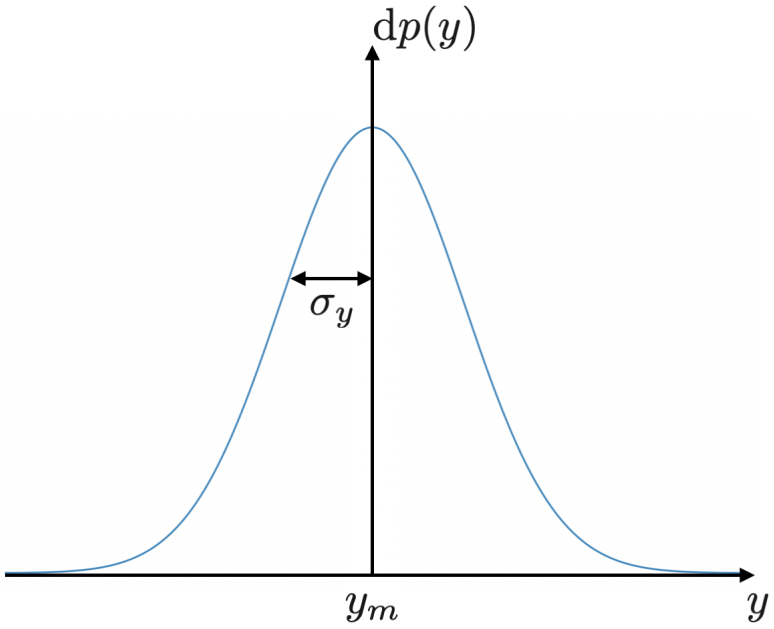



FIGURE 2.4 – Gaussian distribution for an internal variable

and are therefore very small for any macroscopic system. As we see, this is valid whether the variable is intensive or extensive. To be more specific, in the Taylor expansion of $s(y)$ in the vicinity of its most probable value, i.e, $y - y_m \sim \sigma_y$, the successive terms are of order

$$(y-y_m)^n \frac{\partial^n s}{\partial y^n}(y_m) \propto \sigma_y^n \frac{\partial^n s}{\partial y^n}(y_m) \propto N^{n\alpha - n/2} \frac{N}{N^{n\alpha}} = N^{1-n/2} \tag{2.17}$$

which shows that the development is fast converging and indeed there is very little chance for $y$ to deviate significantly from $y_m$. Note that this reasoning may be at fault when the maximum entropy is reached at the boundary of the domain of variations of $y$ (see the example of the Joule & Gay-Lussac expansion below).

### 2.3.2 Relaxation of an isolated system

In this section, we show why the entropy increases necessarily in the evolution of an isolated system towards equilibrium. To fix ideas, consider a situation where an isolated box is initially separated in two compartments by an infinitely thin, insulated wall that is blocked in place. In one compartment we have a gas filling the volume $V$ that is accessible to it. The number of particles is $N$ and the energy of the gas is $E$. In the other compartment (volume $V_0$) is a vacuum (no particles, zero energy). At $t = 0$ we remove, with negligible work, the peg that held the wall in place and the latter is now able to freely move (no friction). This process is called a **Joule & Gay-Lussac expansion**. The gas will push on the wall until it fills the whole volume $V + V_0$. During this process, the energy $E$ and number of particles $N$ remain fixed, but the volume occupied by the gas is an internal variable $V'$ that starts at $V$ and will eventually end up at $V + V_0$. More generally, we may consider what the effect of such a **relaxation** of an internal variable is on the entropy of an isolated system, that is of relaxing a constraint on a variable $X$ (here the volume $V$). We start from $\omega(E, x; \mathrm{d}x) = u(E, x)\mathrm{d}x$ which is the number of accessible states for which $X$ takes values in $[x, x + \mathrm{d}x]$. We note that $u \geqslant 0$. The total number of states for which $X$ lies between $x_A$ and $x_B \geqslant x_A$ is then

$$\Omega(E, x_A, x_B) = \int_{x_A}^{x_B} u(E, x)\mathrm{d}x \tag{2.18}$$

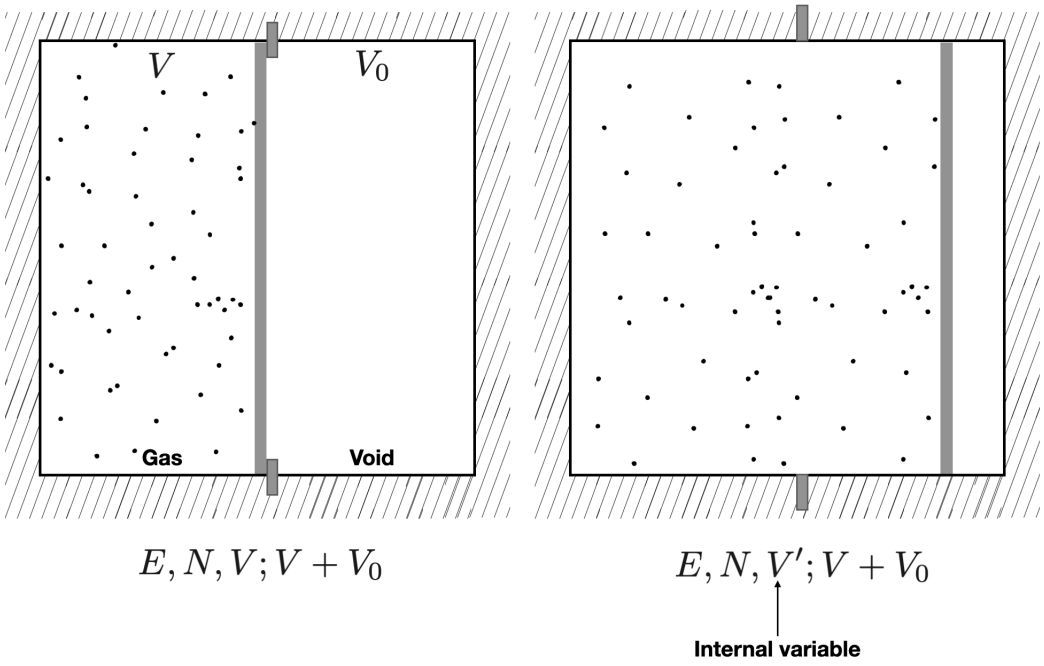



FIGURE 2.5 – Joule & Gay-Lussac expansion

By relaxing a constraint, we allow the system to occupy states covering a broader range of values [1] for $X$, i.e., $[x'_A, x'_B]$ with $x'_A \leqslant x_A$ and $x'_B \geqslant x_B$. Therefore

$$\Omega' = \Omega(E, x'_A, x'_B) = \int_{x'_A}^{x'_B} u(E,x)\mathrm{d}x \geqslant \int_{x_A}^{x_B} u(E,x)\mathrm{d}x = \Omega(E, x_A, x_B) = \Omega \tag{2.19}$$

This may be represented schematically as in Fig. 2.6. The same goes for the logarithm of $\Omega$, that is the entropy, which can thus only increase in the evolution of an isolated system, $S' \geqslant S$. This is most generally written in the following form

$$\mathrm{d}S \geqslant 0 \tag{2.20}$$

## 2.4 Equilibrium between two sub-systems

### 2.4.1 Temperature and thermal equilibrium

#### Equilibrium condition

Consider an isolated system at equilibrium. Its microcanonical entropy $S(E, V, N, x, \ldots)$ may be used to define the **microcanonical temperature** $T$ from the relation

1. In our example, the volume occupied by the gas could be anything between 0 and $V$ initially, and anything between 0 and $V + V_0$ in the end.

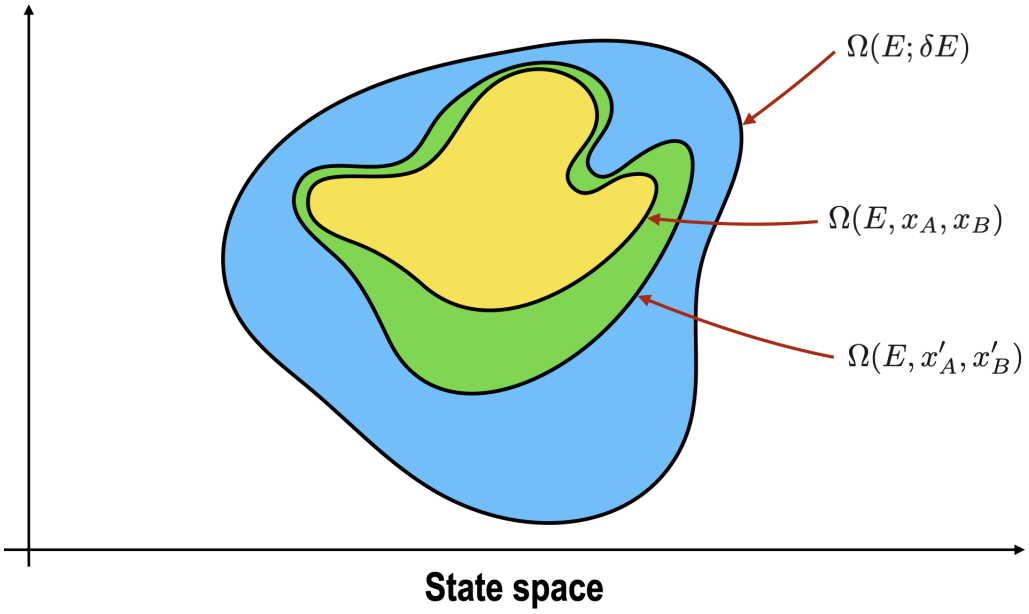



FIGURE 2.6 – The relaxation of a constraint on an internal variable $X$, from $[x_A, x_B]$ to $[x'_A, x'_B]$ with $x'_A \leqslant x_A$ and $x'_B \geqslant x_B$ leads to an increase of the number of compatible microstates, from the yellow subset to the green one.

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{V,N,x,\ldots} \tag{2.21}$$

It may be shown that in all physically-sound cases, the temperature is positive, $T > 0$. We discuss this later on in this chapter.

Now consider two systems $\mathcal{S}_1$ and $\mathcal{S}_2$ that are both isolated, with respective parameters $(E_1, V_1, N_1)$ and $(E_2, V_2, N_2)$, and **weakly coupled**. This means that the energy of the union is

$$E(\mathcal{S}_1 \cup \mathcal{S}_2) = E_1 + E_2 + E_{12} \approx E_1 + E_2 \tag{2.22}$$

i.e., we neglect the interaction energy $E_{12}$ between the two systems. This may of course not be possible in the case of long-range interactions such as gravitation in astrophysical systems. The conditions for such a weak coupling hypothesis are that forces are short-range [2] and that the systems are reasonably large in all dimensions (i.e., they are not sheet-like for instance, in which case surface effects may occur).

In this case, the number of accessible states for the union of the two systems is simply the product of the numbers of accessible states for each system separately, since we can choose one state for each system independently from the other. We thus have

$$\Omega_{\mathcal{S}_1 \cup \mathcal{S}_2} = \Omega_{\mathcal{S}_1} \Omega_{\mathcal{S}_2} \tag{2.23}$$

and the Boltzmann formula leads to showing that **entropy is extensive**

2. Coulomb interactions are long-range, and stronger in magnitude than gravitation, but they are effectively cut short by screening effects (Debye length).

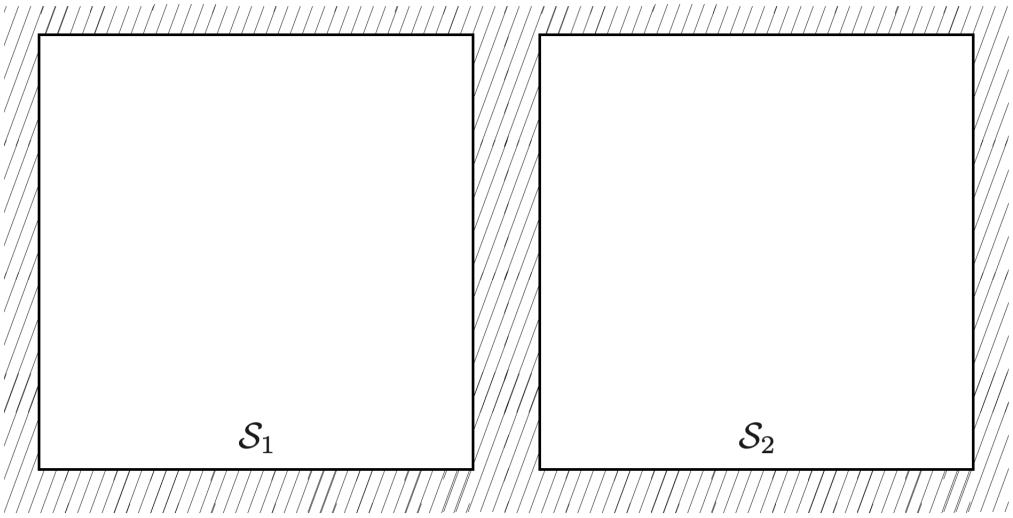



FIGURE 2.7 – Two sub-systems of an isolated system

$$S_{\mathcal{S}_1 \cup \mathcal{S}_2} = S_1 + S_2 \tag{2.24}$$

Now consider that we relax the constraint on the adiabaticity of the wall separating the two systems. The volumes $V_1$ and $V_2$, and the numbers of particles $N_1$ and $N_2$ remain fixed external parameters, as well as the total energy $E = E_1 + E_2$. However, the energy of each system now is an internal variable. Since the sum of the two is constant, we may focus on just one of them, say $E_1$. The most probable value $E_{1,m}$ for this variable is the one that maximizes the entropy

$$S(E_1; V_1, N_1, V_2, N_2, E) = S_1(E_1; V_1, N_1) + S_2(E_2; V_2, N_2) \tag{2.25}$$

taken as a function of $E_1$. Writing explicitly the condition for finding the maximum, we have

$$\frac{\partial S}{\partial E_1} = \frac{\partial S_1}{\partial E_1} + \frac{\partial S_2}{\partial E_1} = \frac{\partial S_1}{\partial E_1} + \frac{\partial S_2}{\partial E_2}\left(\frac{\partial E_2}{\partial E_1}\right)_E = \frac{\partial S_1}{\partial E_1} - \frac{\partial S_2}{\partial E_2} = \frac{1}{T_1} - \frac{1}{T_2} = 0 \tag{2.26}$$

The equilibrium condition is therefore that **the temperatures are equal**, $T_1 = T_2$, or more explicitly,

$$T_1\left(E_{1,m}, V_1, N_1\right) = T_2\left(E - E_{1,m}, V_2, N_2\right) \tag{2.27}$$

This equation may be solved to determine the value $E_{1,m}$ of $E_1$ at equilibrium. It also justifies that the quantity $T$ defined this way is indeed a temperature : it is equal for the two subsystems that are allowed to exchange heat. We note that temperature, just like entropy, is a quantity that is statistical in nature. It requires a macroscopic system to acquire its meaning.

#### Stability condition

We may also derive a stability condition by explicitly writing that we seek a maximum (the temperature equality only made use of the fact that $S$ reached an extremum), i.e.,

$$\frac{\partial^2 S}{\partial E_1^2} = \frac{\partial^2 S_1}{\partial E_1^2} + \frac{\partial^2 S_2}{\partial E_2^2} < 0 \tag{2.28}$$

Assume that the two systems are identical, so that, by symmetry, their energies at equilibrium are identical, $E_1 = E_2$, and so are their entropies. The above relation gives

$$2\frac{\partial^2 S'}{\partial E'^2} < 0 \tag{2.29}$$

where $E'$ and $S'$ are the energy and entropy of each of the systems. Since this can be done for any splitting of any system, provided that the sub-systems remain macroscopic and that their couplings remain weak, we conclude that the stability condition above simply reads

$$\frac{\partial^2 S}{\partial E^2} < 0 \tag{2.30}$$

We call this a stability condition because if we start from the equilibrium and transfer some energy $\delta E$ from one system to the other, the change in entropy is

$$\delta S = \frac{\delta E^2}{2}\left[\frac{\partial^2 S_1}{\partial E_1^2} + \frac{\partial^2 S_2}{\partial E_2^2}\right] < 0 \tag{2.31}$$

and so the entropy ends up smaller in that displaced state, so that the natural evolution of the system will bring it back to the initial state.

Now this second derivative of entropy with respect to energy may also be written as

$$\frac{\partial^2 S}{\partial E^2} = \frac{\partial}{\partial E}\left(\frac{\partial S}{\partial E}\right) = \frac{\partial}{\partial E}\left(\frac{1}{T}\right) = -\frac{1}{T^2}\frac{\partial T}{\partial E} \tag{2.32}$$

so the stability condition is equivalent to the positivity of the **heat capacity at constant volume**

$$C_v = \left(\frac{\partial E}{\partial T}\right)_{V,N} > 0 \tag{2.33}$$

The direction of heat exchanges is such that heat flows from the "hot" system to the "cold" one. This may be found by considering graphs of $T_1$ and $T_2$ as a function of $E_1$. As noted, the positivity of the heat capacity means that $T_1$ is an increasing function of $E_1$, and $T_2$ is also an increasing function of $E_2$. However, since $E_2 = E - E_1$, $T_2$ is a decreasing function of $E_1$. The equilibrium point is where the two curves meet. If initially $T_2^{(i)} > T_1^{(i)}$ we are to the left of this point, $E_1$ increases from its initial value to the final one. If on the other hand we initially have $T_2^{(i)} < T_1^{(i)}$ we are to the right of the equilibrium point and $E_1$ will decrease. We note that the final temperature is always in between the initial temperatures of the two systems.

### 2.4.2 Generalization - Pressure and chemical potential

By definition, the **pressure** $P$ and **chemical potential** $\mu$ of an isolated system are determined from other derivatives of the microcanonical entropy

$$\frac{P}{T} = \left(\frac{\partial S}{\partial V}\right)_{E,N,x,\ldots} \qquad \frac{\mu}{T} = -\left(\frac{\partial S}{\partial N}\right)_{E,V,x,\ldots} \tag{2.34}$$

It is easy to verify that the pressure and chemical potential thus defined have the correct dimensions (energy per unit volume and energy per particle, respectively). The identification of that pressure with the one defined from kinetic theory is not immediate, but is ensured by the recovery of the ideal gas equation of state, as we shall see later on. To ensure stability, pressure at equilibrium is necessarily

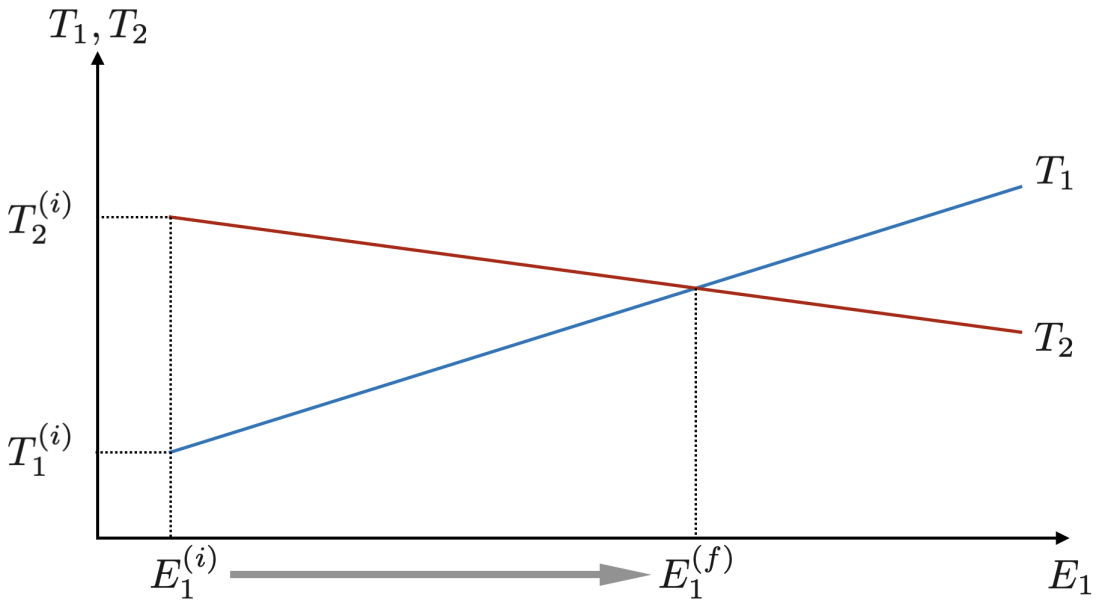



FIGURE 2.8 – Temperatures vs. energy of the first sub-system

positive, $P > 0$, but the chemical potential may be either positive or negative, depending on the system under study.

We note that for any other external parameter $x$, it is possible to define a **conjugate force** $X$ by

$$\frac{X}{T} = \left(\frac{\partial S}{\partial x}\right)_{E,V,N,\ldots} \tag{2.35}$$

To determine the equilibrium conditions, consider an isolated system $\mathcal{S}$ that is made up of two sub-

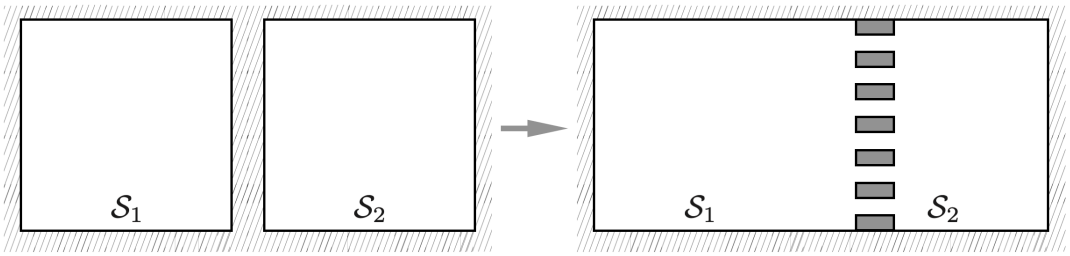



FIGURE 2.9 – Allowing two subsystems to exchange volume, particles, and energy

systems $\mathcal{S}_1$ and $\mathcal{S}_2$ that are initially completely isolated from one another, and imagine that at $t = 0$ we allow the wall separating the two sub-systems to move and let through both particles and heat. This means relaxing all constraints on volume, particle number, and energy for each sub-system, provided the sums $E = E_1 + E_2$, $V = V_1 + V_2$, and $N = N_1 + N_2$ remain constant since the complete system is isolated. The internal variables are then $E_1$, $V_1$, and $N_1$. The entropy of the system is considered as a function of these variables

$$S(E_1, V_1, N_1) = S_1(E_1, V_1, N_1) + S_2(E - E_1, V - V_1, N - N_1) \tag{2.36}$$

and at equilibrium the entropy is maximum with respect to these variables, so

$$\frac{\partial S}{\partial E_1} = 0 \qquad \frac{\partial S}{\partial V_1} = 0 \qquad \frac{\partial S}{\partial N_1} = 0 \tag{2.37}$$

leading, as we did earlier for temperature, to the following equalities

$$T_1 = T_2 \qquad \frac{P_1}{T_1} = \frac{P_2}{T_2} \qquad \frac{\mu_1}{T_1} = \frac{\mu_2}{T_2} \tag{2.38}$$

and therefore to the equality of temperatures, pressures, and chemical potentials

$$T_1 = T_2 \qquad P_1 = P_2 \qquad \mu_1 = \mu_2 \tag{2.39}$$

These allow in turn to determine the values of the maximally-probable energy $E_{1,m}$, volume $V_{1,m}$, and number of particles $N_{1,m}$. The direction of exchanges of volume and particles may be derived quite similarly to what was done for heat exchanges earlier, by writing out the first order entropy variation away from equilibrium and assuming that thermal equilibrium is already achieved. By allowing only volume exchanges or particle exchanges, it is straightforward to show that, when returning to equilibrium, the system with greater pressure gains volume and the system with the largest chemical potential loses particles. We leave it to the reader to show this.

Note that if the two systems are separated by an adiabatic piston, there is nonetheless some energy exchange between the two subsystems due to the work of pressure forces. It is therefore quite challenging to imagine an exchange of volume at constant energy. The same is true of particle exchanges through a porous wall, since particles carry energy along with them.

## 2.5 Examples

### 2.5.1 The perfect paramagnetic crystal

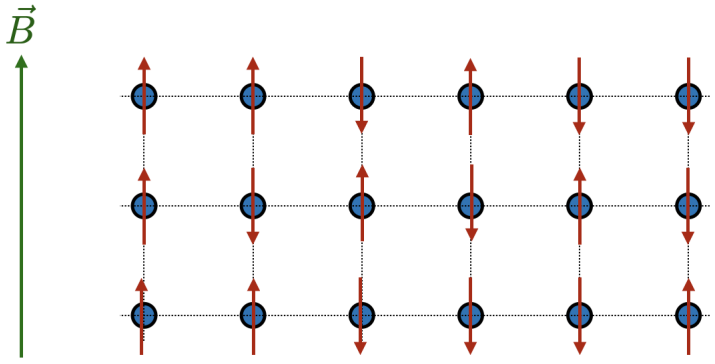



FIGURE 2.10 – The perfect paramagnetic crystal

#### Entropy and canonical temperature

Reprise the system discussed in the previous chapter, where $N$ independent atoms at fixed positions are assigned spins, with $s = 1/2$, that can be either pointing "up" or "down" in an ambiant magnetic

field $\vec{B} = B\vec{e}_z$. The magnetic moment $\vec{\mu}$ may take the value $\mu\vec{e}_z$ ("up" state) or $-\mu\vec{e}_z$ ("down" state). We write $n_+$ for the number of atoms in the "up" state and $n_- = N - n_+$ for the number of atoms in the "down" state. The associated energies are $\epsilon_+ = -\mu B$ and $\epsilon_- = \mu B$, respectively. The energy of the macrostate is a function of one parameter, say $n_+$, with

$$E(n_+) = \epsilon_+ n_+ + \epsilon_- n_- = \mu B\,(N - 2n_+) \tag{2.40}$$

and the corresponding number of accessible states is

$$\Omega\,(n_+) = C_N^{n_+} = \frac{N!}{n_+!\,(N - n_+)!} \tag{2.41}$$

Since the system is isolated, $E$ and $N$ are fixed, so $n_+$ is fixed as well. Using the Stirling formula for the factorials, we have the entropy

$$S \approx k_B\,[N \ln N - n_+ \ln n_+ - (N - n_+) \ln\,(N - n_+)] \tag{2.42}$$

From this we may compute the temperature, i.e.,

$$\frac{1}{T} = \frac{\partial S}{\partial E} = \frac{\partial S}{\partial n_+}\frac{\partial n_+}{\partial E} = [-\ln n_+ + \ln\,(N - n_+)] \times \left(-\frac{k_B}{2\mu B}\right) \tag{2.43}$$

and from there derive the number of "up" state and "down" state atoms [3]

$$n_+ = N\frac{\exp\left(\dfrac{\mu B}{k_B T}\right)}{2\cosh\left(\dfrac{\mu B}{k_B T}\right)} \qquad n_- = N\frac{\exp\left(-\dfrac{\mu B}{k_B T}\right)}{2\cosh\left(\dfrac{\mu B}{k_B T}\right)} \tag{2.47}$$

We note that the temperature is negative if $n_+ < N/2$, which corresponds to $E > 0$,

$$T = -\frac{2\mu B}{k_B}\frac{1}{\ln\left(\dfrac{N}{n_+} - 1\right)} < 0 \tag{2.48}$$

In fact, $S$ is an increasing function of $E$ for negative energy, but a decreasing one for positive energy, as shown in Fig. 2.11. In this latter case, we have a **population inversion**, in that there are more spins with a higher energy $\epsilon_-$ than there are spins with a lower energy $\epsilon_+$. The occurrence of such negative temperatures is a necessary consequence of the finite number of states (here, just two) for each particle.

3. The derivation starts by rewriting the equation defining $T$ as

$$\ln\left(\frac{n_+}{N - n_+}\right) = \frac{2\mu B}{k_B T} \tag{2.44}$$

so that, isolating $n_+$ we have

$$n_+ = N\frac{\exp\left(\dfrac{2\mu B}{k_B T}\right)}{1 + \exp\left(\dfrac{2\mu B}{k_B T}\right)} = N\frac{\exp\left(\dfrac{\mu B}{k_B T}\right)}{\exp\left(-\dfrac{\mu B}{k_B T}\right) + \exp\left(\dfrac{\mu B}{k_B T}\right)} = N\frac{\exp\left(\dfrac{\mu B}{k_B T}\right)}{2\cosh\left(\dfrac{\mu B}{k_B T}\right)} \tag{2.45}$$

The computation of $n_- = N - n_+$ is straightforward and leads to

$$n_- = N\left[1 - \frac{\exp\left(\dfrac{\mu B}{k_B T}\right)}{\exp\left(-\dfrac{\mu B}{k_B T}\right) + \exp\left(\dfrac{\mu B}{k_B T}\right)}\right] = N\frac{\exp\left(-\dfrac{\mu B}{k_B T}\right)}{\exp\left(-\dfrac{\mu B}{k_B T}\right) + \exp\left(\dfrac{\mu B}{k_B T}\right)} = N\frac{\exp\left(-\dfrac{\mu B}{k_B T}\right)}{2\cosh\left(\dfrac{\mu B}{k_B T}\right)} \tag{2.46}$$

Indeed, this means that the energy of the system not only has a lower bound, but also an upper bound, with finite entropies both, so that $S(E)$ necessarily presents a maximum in between, and thus a range of energies for which the entropy decreases as energy increases. No such thing happens in the case of an infinite number of individual states. In real physical systems, particles have kinetic energy that can increase indefinitely, so negative temperatures only apply to specific degrees of freedom (here, the particles' spins). To make this a clear distinction, one would speak of a **spin temperature**.

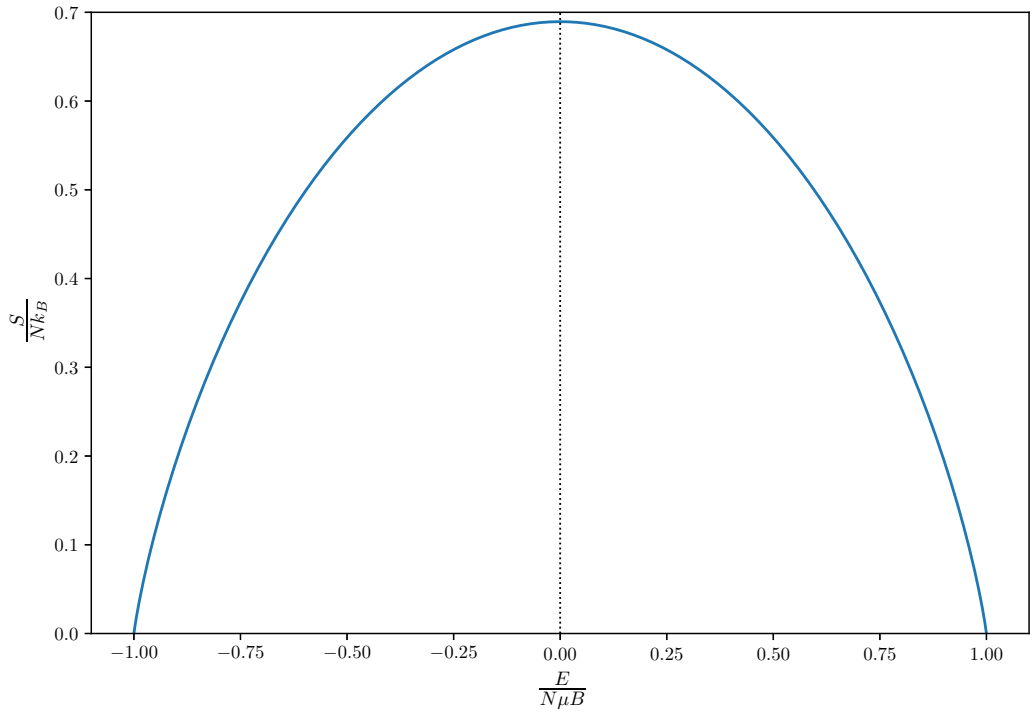



FIGURE 2.11 – Entropy of the perfect paramagnetic crystal as a function of total energy $E$.

### Magnetization and magnetic susceptibility

The **magnetization** $\mathcal{M}$ of the system is defined as the mean magnetic moment per unit volume,

$$\mathcal{M} = \frac{(\mu n_+ - \mu n_-)}{V} = \frac{\mu N}{V} \tanh\left(\frac{\mu B}{k_B T}\right) \tag{2.49}$$

At low temperatures, this leads to $\mu N/V$ which means that all magnetic moments are aligned parallel to the applied field. At high temperatures, $\mathcal{M}$ goes to zero because magnetic moments can thermally flip easily from an "up" to a "down" state and vice-versa, so that on average there are as many up spins as there are down spins. The **magnetic susceptibility** $\chi$ is the coefficient of proportionality between the magnetization and the applied magnetic field at low field, i.e.,

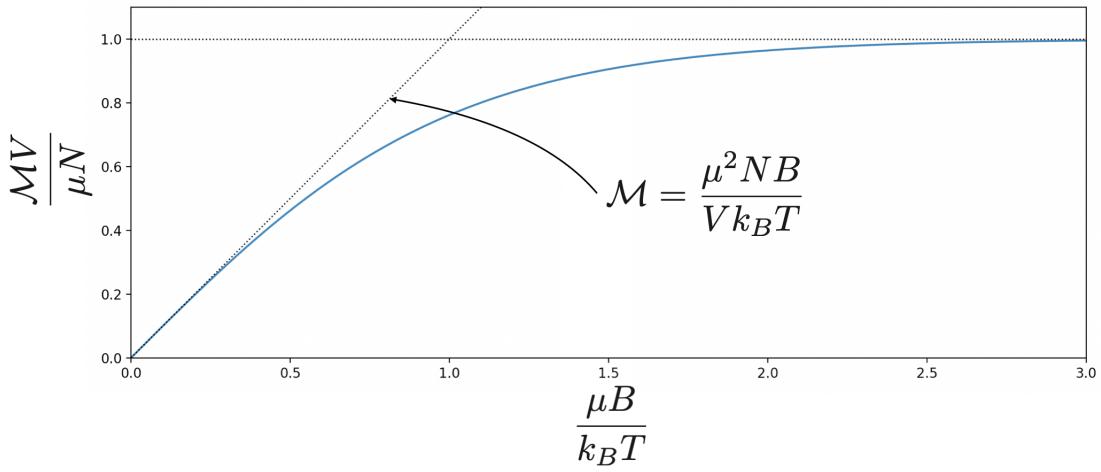



FIGURE 2.12 – Magnetization of the perfect paramagnetic crystal.

$$\chi = \lim_{B\to 0} \frac{\mathcal{M}}{B} = \frac{\mu^2 N}{Vk_BT} = \frac{\mathcal{C}}{T} \tag{2.50}$$

This is the **Curie law**. We will return to this system in Chapter 9.

### 2.5.2 The ideal gas

#### Equation of state

Consider an ideal gas of $N$ monoatomic particles enclosed within a volume $V$, and with total energy $E$. This total energy is only made up of the kinetic energy of the particles, since the gas is ideal (no interactions). As already computed, the density of states is then

$$\rho_N(E) = A_N V^N E^{3N/2-1} \tag{2.51}$$

and the microcanonical entropy is then

$$S = k_B \ln \rho_N \approx Nk_B \left( \ln V + \frac{3}{2} \ln E \right) + k_B \ln A_N \tag{2.52}$$

Remembering the expression of the constant $A_N$, we note that

$$k_B \ln A_N = k_B \ln \left( \frac{3N}{2} \right) + \frac{3Nk_B}{2} \ln \left( \frac{m}{2\pi\hbar^2} \right) - k_B \ln \Gamma \left( \frac{3N}{2} + 1 \right) \tag{2.53}$$

Using Stirling's formula for the last term and neglecting terms that are of order $\ln N$ compared to those of order $N$, we have, after some simple manipulations left to the reader as an exercise,

$$S \approx Nk_B \left[ \ln V + \frac{3}{2} \ln \left( \frac{E}{N} \right) + \frac{3}{2} \ln \left( \frac{m}{3\pi\hbar^2} \right) + \frac{3}{2} \right] = Nk_B \left[ \ln V + \frac{3}{2} \ln \left( \frac{E}{N} \right) + \xi \right] \tag{2.54}$$

The microcanonical temperature is computed from the derivative of $S$ with respect to $E$,

$$\frac{1}{T} = \frac{\partial S}{\partial E} = \frac{3Nk_B}{2E} \tag{2.55}$$

which leads to the well-known expression of the system's (internal) energy as a function of temperature

$$E = \frac{3}{2} N k_B T \tag{2.56}$$

and to the heat capacity at constant volume

$$C_v = \left( \frac{\partial E}{\partial T} \right)_V = \frac{3}{2} N k_B \tag{2.57}$$

The equation of state is also obtainable from the definition of the pressure (left as an exercise)

$$PV = N k_B T \tag{2.58}$$

### Ideal mixture of ideal gases

Let us consider a mixture of two ideal gases, such that there are no more interactions between particles of different types than there are between particles of the same gas. This is an **ideal mixture**. In the volume $V$, there are $N_1$ particles of the first gas, with total energy $E_1$, and $N_2$ particles of the second gas, with total energy $E_2$. The mixture being ideal means that the total energy is simply $E = E_1 + E_2$, and this is constant if the whole system is isolated, as we assume. Note that $E_1$ and $E_2$ are internal variables, since energy exchanges between the two gases are nevertheless possible. Then the density of states of the mixture is

$$\rho(E) \propto V^{N_1+N_2} E^{3(N_1+N_2)/2-1} \tag{2.59}$$

a formula that we leave to the reader to obtain, from the density of states of an ideal gas of $N$ particles given in (1.73). From there, we obtain the microcanonical entropy

$$S = N_1 k_B \left[ \xi_1 + \ln V \right] + N_2 k_B \left[ \xi_2 + \ln V \right] + \frac{3}{2} (N_1 + N_2) k_B \ln \left( \frac{E}{N_1 + N_2} \right) \tag{2.60}$$

where $\xi_1$ and $\xi_2$ are constants depending only on the masses $m_1$ and $m_2$ of the particles. Further on, we obtain the temperature and pressure,

$$E = \frac{3}{2} (N_1 + N_2) k_B T \qquad\qquad PV = (N_1 + N_2) k_B T \tag{2.61}$$

which are the same as for a simple ideal gas, replacing $N$ by the total number of particles. From this we define the **partial pressures**

$$P_1 = \frac{N_1 k_B T}{V} \qquad\qquad P_2 = \frac{N_2 k_B T}{V} \tag{2.62}$$

### The Gibbs paradox

It may not have escaped the reader that the expression of the entropy is not extensive. This is a problem that is best underlined by considering the evolution of an ideal mixture of two gases that are initially separated in two distinct compartments. Removing the wall, particles of the two gases mingle, and the entropy in the final state is larger than in the initial state. The variation $\Delta S > 0$, which we leave to the reader to compute, is called the **entropy of mixture**. Now imagine these two gases are in fact identical. The macrostates before and after removing the wall are then identical, so we should have $\Delta S = 0$. This is the **Gibbs paradox** that can only be solved by taking into account the fundamental indistinguishability of the particles, as we shall see.

## 2.6 The case of an isolated system in macroscopic motion

### 2.6.1 Conditions for statistical equilibrium

For completeness, we discuss here briefly the important case of an isolated system $\mathcal{S}$ undergoing macroscopic motion. The conditions for it to be in statistical equilibrium may be derived by dividing it into infinitesimal (but still macroscopic) volume elements, labelled with an index $\alpha$. At any given time, each of these contains a certain number of particles $N_\alpha$, encloses a given total energy $E_\alpha$, is located at a certain position $\vec{r}_\alpha$ and is imprinted with a momentum $\vec{P}_\alpha$. Over the whole system, the assumption that it is isolated implies conservation of the total energy $E$, number of particles $N$, momentum $\vec{P}$, and angular momentum $\vec{L}$. Now, as time goes by, the different volume elements may exchange particles, energy, momentum, and angular momentum. The condition for statistical equilibrium of the system implies that the entropy must be maximum with respect to the various variables $E_\alpha$, $N_\alpha$, and $\vec{P}_\alpha$. This can be expressed together with the conservation of $E$, $N$, $\vec{P}$, and $\vec{L}$ using the method of **Lagrange multipliers** already seen in the previous chapter. We shall not give the full derivation here, but the reader is encouraged to do this as an exercise. The physical interpretation of the results of this procedure is that for the system to be in statistical equilibrium requires :

— the temperature to be the same everywhere ;
— the system to move in solid body rotation, without deformation, about a principal axis of inertia ;
— the sum of the chemical potential and the centrifugal potential to be the same everywhere.

### 2.6.2 Consequences

#### Rotation about the axis of maximum moment of inertia

Note that in the center-of-mass frame, the entropy is a function of $E$, $N$, and $\vec{L}$ only, and reads

$$S = S\left(E - \frac{\vec{L}^2}{2I}, N\right) \tag{2.63}$$

where $I$ is the **moment of inertia** of the axis about which the system rotates. From this, it is easy to see that, in statistical equilibrium, for a given energy, number of particles, and angular momentum, rotation occurs about the axis for which $I$ is maximum.

#### Positivity of the temperature

For simplicity, consider a case where the system is studied in its center-of-mass frame, and does not rotate. Its entropy reads as the sum of the entropies of the various volume elements,

$$S = \sum_\alpha S_\alpha\left(E_\alpha - \frac{\vec{P}_\alpha^2}{2mN_\alpha}\right) \tag{2.64}$$

subject to the fact that the sum of all momenta is constant, and null in this case. Should the temperature (which must be the same for all volume elements), be negative, the various $S_\alpha$ would be decreasing functions of the corresponding internal energies. Consequently, by increasing the different momenta, and therefore decreasing the internal energies, the $S_\alpha$ would increase, which would be the natural evolution of the system. It is not difficult to imagine that this means an explosion or implosion of the system. Stability of the equilibrium therefore imposes that $T > 0$.

# 3 Canonical ensemble

In this chapter, we discuss the **canonical distribution**, which gives the probabilities of the various microstates for a system that is in contact with a **thermostat**, that is a (much larger) system imposing its temperature to the system. This means that the system is free to exchange energy with the thermostat, but the other parameters (volume and number of particles) are fixed.

## 3.1 Canonical distribution

### 3.1.1 Thermostat and canonical temperature

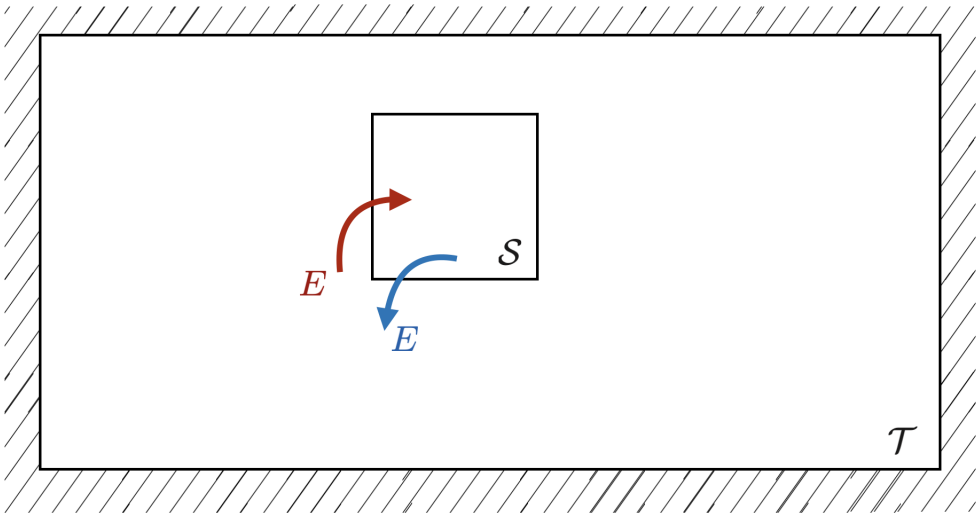



FIGURE 3.1 – A system $\mathcal{S}$ in contact with a thermostat $\mathcal{T}$

Consider a system $\mathcal{S}$ that is in purely thermal contact with a much larger system $\mathcal{T}$, which is called a thermostat (or **heat bath**). The number of particles $N$ and the volume $V$ of $\mathcal{S}$ are fixed external parameters, while its energy $E$ is now an internal variable. We assume that the coupling between $\mathcal{S}$ and $\mathcal{T}$ is weak and that the total system $\mathcal{S} \cup \mathcal{T}$ is isolated so that the total energy $E_{\mathcal{S}\cup\mathcal{T}} = E + E_{\mathcal{T}}$ is conserved and is therefore also an external parameter, whose value we will write $E_0$. The fact that $\mathcal{T}$

is much larger than $\mathcal{S}$ implies that

$$E \ll E_{\mathcal{S}\cup\mathcal{T}} = E_0 \approx E_{\mathcal{T}} \tag{3.1}$$

As a result, energy exchanges have little effect on the thermostat, so the microstates of $\mathcal{T}$ are virtually independent of the state of $\mathcal{S}$. We may consider the microcanonical entropy of $\mathcal{T}$ as $S_{\mathcal{T}}(E_0 - E)$ and derive the microcanonical temperature of the thermostat as

$$T_{\mathcal{T}} = \left(\frac{\partial S_{\mathcal{T}}}{\partial E_{\mathcal{T}}}\right)^{-1} = T_{\mathcal{T}}(E_0 - E) \tag{3.2}$$

The thermostat would be perfect if its temperature did not depend on the energy content of the system $\mathcal{S}$ with which it is in contact. A Taylor expansion helps to quantify this through

$$\frac{1}{T_{\mathcal{T}}}(E_0 - E) = \frac{1}{T_{\mathcal{T}}}(E_0) - E\frac{\partial^2 S_{\mathcal{T}}}{\partial E_{\mathcal{T}}^2}(E_0) \tag{3.3}$$

In this expression, the first term on the right-hand side is of the order $\sim N_{\mathcal{T}}/N_{\mathcal{T}} = 1$, while the second term is of the order $\sim NN_{\mathcal{T}}/N_{\mathcal{T}}^2 = N/N_{\mathcal{T}} \ll 1$. In that limit, $T_{\mathcal{T}}$ becomes independent of $E$, and is therefore an external parameter for the system $\mathcal{S}$. We will write it $T$ in this case.

To summarize, in the microcanonical framework, $E$, $V$, and $N$ are external parameters, and we derive the temperature $T$ as an internal variable. In the canonical framework, the external parameters are $T$, $V$, and $N$, and the energy $E$ is now an internal variable.

### 3.1.2 Probabilities of microstates of $\mathcal{S}$

We write $|\ell\rangle$ for the microstates of $\mathcal{S}$, and $|L\rangle$ for the microstates of $\mathcal{T}$. These have respective energies $E_\ell$ and $E_L$, and the total energy is $E_0 = E_\ell + E_L$ in the weak coupling regime. The states of the full system are noted $|\ell, L\rangle$, and their probabilities $p_{\ell,L}$ are determined by the microcanonical distribution, since the full system is isolated,

$$p_{\ell,L} = \frac{1}{\Omega_{\mathcal{S}\cup\mathcal{T}}(E_0)} \tag{3.4}$$

For $\mathcal{S}$ to be in a state $|\ell\rangle$ whose energy is $E_\ell$, the thermostat must be in a state $|L\rangle$ such that its energy is

$$E_L = E_0 - E_\ell \tag{3.5}$$

within the experimental uncertainty $\delta E$. There are basically $\Omega_{\mathcal{T}}(E_0 - E_\ell)$ such states, so

$$p_\ell = \frac{\Omega_{\mathcal{T}}(E_0 - E_\ell)}{\Omega_{\mathcal{S}\cup\mathcal{T}}(E_0)} = C\exp\left[\frac{S_{\mathcal{T}}(E_0 - E_\ell)}{k_B}\right] \tag{3.6}$$

inserting here the microcanonical entropy of the thermostat. Now, since $E_\ell \ll E_0$, we may write this entropy as a Taylor expansion

$$S_{\mathcal{T}}(E_0 - E_\ell) = S_{\mathcal{T}}(E_0) - E_\ell\frac{\partial S_{\mathcal{T}}}{\partial E_{\mathcal{T}}}(E_0) + \ldots \approx S_{\mathcal{T}}(E_0) - \frac{E_\ell}{T} \tag{3.7}$$

and the probabilities of the various microstates of $\mathcal{S}$ are then

$$p_\ell = C'\exp\left(-\frac{E_\ell}{k_B T}\right) \tag{3.8}$$

The normalisation condition states that the sum of these probabilities is unity, which implies that

$$p_\ell = \frac{1}{Z}\exp\left(-\frac{E_\ell}{k_B T}\right) \tag{3.9}$$

where we have introduced the **canonical partition function**

$$Z = \sum_{|\ell\rangle} \exp\left(-\frac{E_\ell}{k_B T}\right) \tag{3.10}$$

These probabilities constitute the **canonical distribution** [1].

All microstates $|\ell\rangle$ are accessible, but their probabilities depending on the energy $E_\ell$, the most populated states are the ones with lowest energy. The ratio between the probabilities of two different states $|\ell\rangle$ and $|\ell'\rangle$ is

$$\frac{p_\ell}{p_{\ell'}} = \exp\left[-\frac{(E_\ell - E_{\ell'})}{k_B T}\right] = \exp\left(-\frac{\Delta E}{k_B T}\right) \tag{3.11}$$

Therefore if the energy difference between the two states is $|\Delta E| \ll k_B T$ then both states are similarly probable, $p_\ell \approx p_{\ell'}$, while if $|\Delta E| \gg k_B T$ then the lower energy state is much more probable than the other one. Orders of magnitude to keep in mind are that $k_B T = 1\,\mathrm{eV}$ corresponds to the very high temperature of $T \simeq 12000\,\mathrm{K}$, and that room temperatures $T \simeq 300\,\mathrm{K}$ correspond to about $25\,\mathrm{meV}$.

### 3.1.3 Distribution of an internal variable

As a typical example of an internal variable, we may consider the case of the energy $E$ of the system $\mathcal{S}$. The probability $p(E)$ for the energy is the product of the probability (3.9) by the number of states whose energy is $E$, i.e., the degeneracy $g(E)$. Between $E$ and $E + \mathrm{d}E$, this degeneracy is $g(E) = \rho(E)\mathrm{d}E$, so the probability is

$$p(E) = w(E)\mathrm{d}E = \frac{1}{Z}\exp\left(-\frac{E}{k_B T}\right)\rho(E)\mathrm{d}E \tag{3.12}$$

introducing the probability density $w(E)$ for the energy. Note that this is possible only because the typical energy difference between two successive levels is much smaller than $k_B T$. The partition function may then be written as an integral over the energy

$$Z = \sum_{|\ell\rangle} \exp\left(-\frac{E_\ell}{k_B T}\right) = \sum_{E_\ell} g(E_\ell)\exp\left(-\frac{E_\ell}{k_B T}\right) = \int \rho(E)\exp\left(-\frac{E}{k_B T}\right)\mathrm{d}E \tag{3.13}$$

Consider the case of a macroscopic system $\mathcal{S}$ at room temperature. The energies involved are typically $1\,\mathrm{J} \gg k_B T$ so apart from the few levels close to the fundamental $E_\ell \approx 0$, the exponential factor in the probability distribution for $E$ in equation (3.12) is negligible and is therefore a fast decreasing function of $E$. However, it is multiplied by a fast increasing function of energy, that is $\rho(E) \propto E^{\alpha N}$. Consequently, $w(E)$ peaks at some finite energy $E_m$ that we mean to determine [2].

To do that, we write the logarithm of the probability density, and seek its maximum

$$\frac{\partial \ln[w(E)]}{\partial E} = \frac{\partial}{\partial E}\left[C - \frac{E}{k_B T} + \ln[\rho(E)]\right] = -\frac{1}{k_B T} + \frac{1}{k_B}\frac{\partial S^*}{\partial E} = 0 \tag{3.14}$$

1. It may be shown that the only two distributions of probabilities that do not depend on the choice of the zero-level energy are the microcanonical and canonical ones. We also note that while the microcanonical approach requires the system to be macroscopic, the canonical approach may be applied to a microscopic system provided the coupling with the thermostat is weak.

2. This reasoning is akin to that leading to the Gamow peak in nuclear reactions.

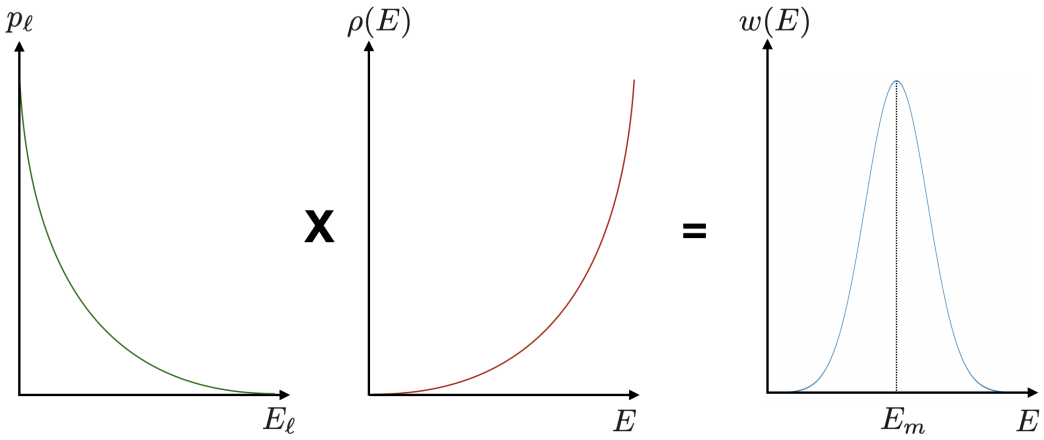



FIGURE 3.2 – Canonical probability, density of states, and distribution of the energy.

where $S^*(E) = k_B \ln[\rho(E)]$ is the microcanonical entropy of the system $\mathcal{S}$. The last term in the equation above leads to the appearance of the microcanonical temperature $T^*$ of $\mathcal{S}$, defined by

$$\frac{1}{T^*} = \frac{\partial S^*}{\partial E} \tag{3.15}$$

so that the most probable value $E_m$ for the energy of $\mathcal{S}$ is the one for which this microcanonical temperature is equal to the canonical temperature imposed by the thermostat

$$T^*(E_m) = T \tag{3.16}$$

Similarly to the previous discussions, we may write the probability distribution of the energy in the vicinity of $E_m$ through a Taylor expansion, leading to a Gaussian distribution

$$w(E) = w(E_m) \exp\left[-\frac{(E-E_m)^2}{2\sigma_E^2}\right] \tag{3.17}$$

where the standard deviation is

$$\sigma_E = \left[-\frac{1}{k_B}\frac{\partial^2 S^*}{\partial E^2}(E_m)\right]^{-1/2} \tag{3.18}$$

## 3.2 Partition function and Helmholtz free energy

### 3.2.1 Definitions

#### Partition function

As already introduced, the **canonical partition function** is defined as

$$Z(T,V,N,\ldots) = \sum_{|\ell\rangle} \exp\left(-\frac{E_\ell}{k_B T}\right) \tag{3.19}$$

or, if the energy may be considered as a continuous variable (in a macroscopic system),

$$Z(T,V,N,\ldots) = \int_0^\infty \rho(E) \exp\left(-\frac{E}{k_B T}\right) \mathrm{d}E \tag{3.20}$$

A very common and useful notation is the following, for the inverse temperature,

$$\beta = \frac{1}{k_B T} \tag{3.21}$$

with which the partition function may be written in the more compact forms

$$Z(\beta, V, N, \ldots) = \sum_{|\ell\rangle} e^{-\beta E_\ell} \qquad Z(\beta, V, N, \ldots) = \int_0^\infty \rho(E) e^{-\beta E} \mathrm{d}E \tag{3.22}$$

We note the useful relation between derivatives with respect to $\beta$ and $T$,

$$\beta \frac{\partial}{\partial \beta} = -T \frac{\partial}{\partial T} \tag{3.23}$$

#### Helmholtz free energy

The **Helmholtz free energy** $F$ is defined from the partition function $Z$ through the relation

$$F(T, V, N, \ldots) = -k_B T \ln Z \tag{3.24}$$

We note how this definition is similar to that of the microcanonical entropy $S = k_B \ln \Omega$ with respect to the number of accessible states $\Omega$. We also note that the definition of $F$ is such that

$$e^{-\beta F} = \sum_{|\ell\rangle} e^{-\beta E_\ell} \tag{3.25}$$

which shows that, in a way, $F$ is some kind of mean energy of the system. However, this is not exact, as we shall see shortly.

### 3.2.2 Properties of the system

#### Mean energy

The mean energy of the system is, by definition,

$$\langle E \rangle = \sum_{|\ell\rangle} E_\ell p_\ell = \frac{1}{Z} \sum_{|\ell\rangle} E_\ell e^{-\beta E_\ell} = -\frac{1}{Z} \frac{\partial}{\partial \beta} \left( \sum_{|\ell\rangle} e^{-\beta E_\ell} \right) = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} \tag{3.26}$$

so that we have finally

$$\langle E \rangle = -\frac{\partial \ln Z}{\partial \beta} \tag{3.27}$$

As exercises[3], we leave it to the reader to show that the mean energy may also be written as follows using the Helmholtz free energy

$$\langle E \rangle = F - T \frac{\partial F}{\partial T} \tag{3.31}$$

3. The demonstrations are as follows, first for the mean energy

$$\langle E \rangle = -\frac{\partial \ln Z}{\partial \beta} = \frac{\partial}{\partial \beta} [\beta F] = F + \beta \frac{\partial F}{\partial \beta} = F - T \frac{\partial F}{\partial T} \tag{3.28}$$

and that the variance of the energy is

$$\sigma_E^2 = \left\langle (E - \langle E \rangle)^2 \right\rangle = \frac{\partial^2 \ln Z}{\partial \beta^2} \tag{3.32}$$

### Heat capacity at constant volume

The heat capacity at constant volume may be computed from the free energy as

$$C_v = \left( \frac{\partial \langle E \rangle}{\partial T} \right)_{V,N,\ldots} = -T \frac{\partial^2 F}{\partial T^2} \tag{3.33}$$

### Distribution of an internal variable

Consider an internal variable $Y$. The probability for it to take the value $y$ is the following sum of $p_\ell$ that extends over the microstates $|\ell\rangle$ in the subset $\mathcal{L}_y$ for which $Y$ takes on the desired value,

$$p(y) = \sum_{|\ell\rangle \in \mathcal{L}_y} p_\ell = \frac{1}{Z} \sum_{|\ell\rangle \in \mathcal{L}_y} e^{-\beta E_\ell} \tag{3.34}$$

This may then be interpreted as $p(y) = z(y)/Z$ where

$$z(y) = \sum_{|\ell\rangle \in \mathcal{L}_y} e^{-\beta E_\ell} \tag{3.35}$$

is the partial partition function, limited to the set of microstates $\mathcal{L}_y$. From this we may define a partial Helmholtz free energy $f(T, V, N; y) = -k_B T \ln [z(y)]$ and the most probable value $y_m$ of $y$ is the one that minimizes this partial free energy. We shall see shortly why the equilibrium state is the one minimizing the Helmholtz free energy.

### Canonical entropy

Similarly to what was done to come up with a microcanonical entropy, it is possible to use the canonical distribution (3.9) to define a **canonical entropy**,

$$S = -k_B \sum_{|\ell\rangle} p_\ell \ln p_\ell \tag{3.36}$$

This may be shown to be related to $F$ and the mean energy $\langle E \rangle$, since

$$S = k_B \sum_{|\ell\rangle} p_\ell \left( \ln Z + \beta E_\ell \right) = k_B \ln Z + \frac{1}{T} \sum_{|\ell\rangle} p_\ell E_\ell = \frac{\langle E \rangle - F}{T} \tag{3.37}$$

and therefore we have the relation already known from thermodynamics

$$F = \langle E \rangle - TS \tag{3.38}$$

---

where the last equality is from (3.23). Then, for the variance, we compute

$$\langle E^2 \rangle = \sum_{|\ell\rangle} E_\ell^2 p_\ell = \frac{1}{Z} \sum_{|\ell\rangle} E_\ell^2 e^{-\beta E_\ell} = \frac{1}{Z} \frac{\partial^2}{\partial \beta^2} \left( \sum_{|\ell\rangle} e^{-\beta E_\ell} \right) = \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2} \tag{3.29}$$

which we use to finalize the computation of

$$\frac{\partial^2 \ln Z}{\partial \beta^2} = \frac{\partial}{\partial \beta} \left[ \frac{1}{Z} \frac{\partial Z}{\partial \beta} \right] = -\frac{1}{Z^2} \left( \frac{\partial Z}{\partial \beta} \right)^2 + \frac{1}{Z} \frac{\partial^2 Z}{\partial \beta^2} = -\langle E \rangle^2 + \langle E^2 \rangle = \sigma_E^2 \tag{3.30}$$

**Canonical pressure and chemical potential**

Much like what was done in the microcanonical framework, we may define a **canonical pressure** and a **canonical chemical potential** from the following relations.

$$P = -\left(\frac{\partial F}{\partial V}\right)_{T,N,\dots} \qquad \mu = \left(\frac{\partial F}{\partial N}\right)_{T,V,\dots} \tag{3.39}$$

For completeness, we note that the last partial derivative is the opposite of the entropy

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N,\dots} \tag{3.40}$$

### 3.2.3 Spontaneous evolution of a system

Consider the system $\mathcal{S}$ in thermal contact with the thermostat $\mathcal{T}$. The complete system $\mathcal{S}\cup\mathcal{T}$ is an isolated system, so if a constraint internal to $\mathcal{S}$ is relaxed, the microcanonical entropy $S_{\mathcal{S}\cup\mathcal{T}}$ can only increase on the path towards the equilibrium, as we saw in the last chapter. Now this entropy is

$$S_{\mathcal{S}\cup\mathcal{T}} = -k_B \sum_{|\ell,L\rangle} p_{\ell,L} \ln p_{\ell,L} \tag{3.41}$$

where the sum extends over all microstates $|\ell, L\rangle$ of the complete system and these states' probabilities are $p_{\ell,L} = p_\ell p_L(E_0 - E_\ell)$, with $E_0 = E_\ell + E_L$ the total energy of the system. From this we have

$$S_{\mathcal{S}\cup\mathcal{T}} = -k_B \sum_{|\ell,L\rangle} p_\ell p_L(E_0 - E_\ell) \left[\ln p_\ell + \ln\left[p_L(E_0 - E_\ell)\right]\right] \tag{3.42}$$

which we may rewrite by separating the sums over the states of $\mathcal{S}$ and the states of $\mathcal{T}$,

$$S_{\mathcal{S}\cup\mathcal{T}} = -k_B \sum_{|\ell\rangle} p_\ell \ln p_\ell \sum_{|L\rangle} p_L(E_0 - E_\ell) - k_B \sum_{|\ell\rangle} p_\ell \sum_{|L\rangle} p_L(E_0 - E_\ell) \ln\left[p_L(E_0 - E_\ell)\right] \tag{3.43}$$

The first sum in the first term is related to the entropy of the system $\mathcal{S}$, the second sum is unity due to the normalization condition for the states of the thermostat compatible with the energy of $\mathcal{S}$, and the second sum in the second term is related to the entropy of the sole thermostat $\mathcal{T}$, so

$$S_{\mathcal{S}\cup\mathcal{T}} = S + \sum_{|\ell\rangle} p_\ell S_{\mathcal{T}}(E_0 - E_\ell) \tag{3.44}$$

In that last sum, the entropy $S_{\mathcal{T}}$ may be developed in the vicinity of $E_0$ since $E_\ell \ll E_0$, so

$$S_{\mathcal{S}\cup\mathcal{T}} = S + \sum_{|\ell\rangle} p_\ell \left[S_{\mathcal{T}}(E_0) - \frac{E_\ell}{T}\right] = S + S_{\mathcal{T}}(E_0) - \frac{\langle E\rangle}{T} = S_{\mathcal{T}}(E_0) - \frac{F}{T} \tag{3.45}$$

Since $S_{\mathcal{T}}(E_0)$ is a constant, the relaxation of an internal constraint leads to an increase of the entropy $S_{\mathcal{S}\cup\mathcal{T}}$ of the complete system, or, equivalently, to a **decrease of the Helmholtz free energy** $F$ of the system $\mathcal{S}$ on the path to equilibrium.

$$\mathrm{d}F \leqslant 0 \tag{3.46}$$

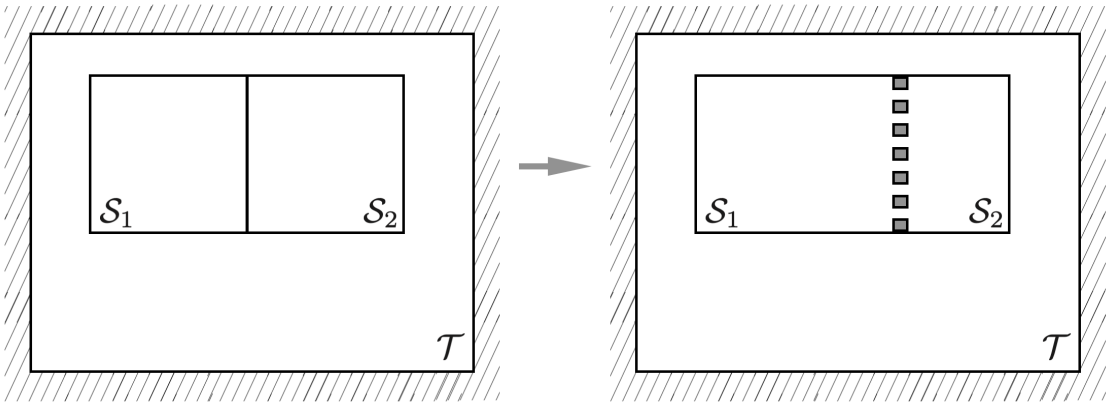



FIGURE 3.3 – Relaxation of constraints in the canonical approach

As an application, we consider the case of a system $\mathcal{S}$ made of two sub-systems $\mathcal{S}_1$ and $\mathcal{S}_2$, separated by a wall that is rendered mobile and permeable to particles at $t = 0$. The quantities $V_1$ and $N_1$ are now internal variables susceptible to change in the process. The volume and number of particles in the other subsystem are $V_2 = V_0 - V_1$ and $N_2 = N_0 - N_1$. The free energy of the system is

$$F(V_0, N_0, T; V_1, N_1) = F_1(T, V_1, N_1) + F_2(T, V_2, N_2) \tag{3.47}$$

where the additivity is left to the reader to demonstrate. At equilibrium, we have the two conditions

$$\frac{\partial F}{\partial V_1} = 0 \qquad \frac{\partial F}{\partial N_1} = 0 \tag{3.48}$$

which lead to the equalities of the canonical pressures and chemical potentials in the two sub-systems,

$$P_1 = P_2 \qquad \mu_1 = \mu_2 \tag{3.49}$$

The stability of the equilibrium requires conditions on the second derivatives of the free energy, so that it is effectively a minimum,

$$\frac{\partial^2 F}{\partial V_1^2} > 0 \qquad \frac{\partial^2 F}{\partial N_1^2} > 0 \tag{3.50}$$

which may be written in terms of derivatives of the chemical potential and of the pressure. Using more general notations, we must have, for the equilibrium to be stable

$$\left(\frac{\partial \mu}{\partial N}\right)_{T,V,\ldots} > 0 \qquad \left(\frac{\partial P}{\partial V}\right)_{T,N,\ldots} < 0 \tag{3.51}$$

## 3.3 The thermodynamic limit

For macroscopic systems, there is an equivalence between the microcanonical and canonical descriptions. That is the **thermodynamic limit**. By macroscopic systems, we consider that both $V \to \infty$ and $N \to \infty$ while the density $N/V$ remains constant.

The reason for this equivalence lies in the fact that in this limit, the fluctuations of an internal variable become negligibly small, and the Gaussian distribution of its possible values extremely peaked

around the mean (also most probable) value. Therefore, there is virtually nothing to distinguish this variable from an external parameter, whose value is fixed by definition.

As an example, in the microcanonical approach, $E$, $V$, and $N$ are fixed and the microcanonical temperature is an internal variable $T^*(E,V,N)$. In the canonical approach, the fixed parameters are $T$, $V$, and $N$, and we computed the mean energy $\langle E\rangle(T,V,N)$ from equation (3.27). If fluctuations of $T^*(E,V,N)$ in the microcanonical approach and of $\langle E(T,V,N)\rangle$ in the canonical approach can be neglected, the relation between the four quantities is the same in both approaches, i.e.,

$$\mathcal{F}(E,V,N,T^*) = \mathcal{F}(\langle E\rangle,V,N,T) = 0 \tag{3.52}$$

This is particularly useful for the resolution of problems, as one can use the statistical description in which computations are easier. We shall see an example later on.

## 3.4 Canonical description of independent particles

### 3.4.1 Setup

Consider $N$ identical particles that are independent, that is, we may neglect the interactions between them. The Hamiltonian of this system then reduces to a sum of terms, each related to one particle only

$$\mathcal{H} = \sum_{i=1}^{N} H_i + V\left(\vec{r}_1,\ldots,\vec{r}_N\right) \approx \sum_{i=1}^{N} H_i \tag{3.53}$$

This system of $N$ particles is considered in the canonical description, so we picture it as being in thermal contact with a thermostat imposing the temperature $T$. The possible states for each individual particle are labelled $|\lambda\rangle$, with energies $\epsilon_\lambda$. The system in its entirety is then described by a state $|\ell\rangle = |\lambda_1,\ldots,\lambda_N\rangle$ with an energy $E_\ell = \epsilon_{\lambda_1} + \ldots + \epsilon_{\lambda_N}$, since there is no interaction energy. The partition function is then

$$Z = \sum_{|\ell\rangle} e^{-\beta E_\ell} \tag{3.54}$$

This may be rewritten in a form that underscores the role played by individual states

$$Z = \sum_{|\lambda_1,\ldots,\lambda_N\rangle} e^{-\beta(\epsilon_{\lambda_1}+\ldots+\epsilon_{\lambda_N})} = \sum_{|\lambda_1,\ldots,\lambda_N\rangle} e^{-\beta\epsilon_{\lambda_1}} \times \ldots \times e^{-\beta\epsilon_{\lambda_N}} \tag{3.55}$$

Now if the summation were over the individual states, it would be usually easy to compute, at least in the case of a perfect gas with no external field, but the summation is over the states of the system of $N$ particles, and that summation depends critically on whether particles are **distinguishable** or not.

### 3.4.2 Distinguishable particles

In the former case, for instance if all particles are assigned at a given spatial site and cannot move significantly away from it, the individual state of particle $i$ is independent of that of particle $j$. In that case, the state $|\ell\rangle = |\lambda_1,\ldots,\lambda_N\rangle$ can be thought of as a "product" $|\ell\rangle = |\lambda_1\rangle \times \ldots \times |\lambda_N\rangle$ and the partition function may be written as such

$$Z = \prod_{i=1}^{N} \sum_{|\lambda_i\rangle} e^{-\beta\epsilon_{\lambda_i}} = \prod_{i=1}^{N} z_i = z^N \tag{3.56}$$

where $z$ is the partition function for each individual particle

$$z = \sum_{|\lambda\rangle} e^{-\beta\epsilon_\lambda} \tag{3.57}$$

As an example, consider the case of a system of two particles $a$ and $b$, each possibly in one of three states, with energies $\epsilon_0$, $\epsilon_1$, and $\epsilon_2$. The partition function reads

$$Z = e^{-\beta(\epsilon_0+\epsilon_0)} + e^{-\beta(\epsilon_0+\epsilon_1)} + e^{-\beta(\epsilon_0+\epsilon_2)} + \ldots + e^{-\beta(\epsilon_2+\epsilon_0)} + e^{-\beta(\epsilon_2+\epsilon_1)} + e^{-\beta(\epsilon_2+\epsilon_2)} \tag{3.58}$$

which does indeed factorize into

$$Z = \left(e^{-\beta\epsilon_0} + e^{-\beta\epsilon_1} + e^{-\beta\epsilon_2}\right)^2 = z^2 \tag{3.59}$$

The free energy of the system is then

$$F = -Nk_BT \ln z \tag{3.60}$$

### 3.4.3 Indistinguishable particles

When particles are not distinguishable, two possibilities arise. Either these particles are **fermions**, with half-integer spin [4] $s$, or they are **bosons** with integer spin. Examples of the former are electrons, for instance, with $s = 1/2$, while $^4\mathrm{He}$ nuclei are bosons. The wave function of fermionic systems must be fully antisymmetric, i.e., it must change sign when two particles are exchanged (positions, momenta, spins), which leads to **Pauli's exclusion principle**, while that of bosonic systems must be fully symmetric.
Now consider how this must affect the counting of possible states in the partition function $Z$. For that consider a system of two particles, $a$ and $b$, each one being possibly in one of three states $|\lambda_1\rangle$, $|\lambda_2\rangle$, and $|\lambda_3\rangle$. Looking at Fig. 3.4, that shows the possible states of the complete system, the indistinguishability of the two particles implies that states $|A\rangle = |\lambda_1, \lambda_2\rangle$ and $|B\rangle = |\lambda_2, \lambda_1\rangle$ are actually the same state, and should be counted only once. The case of state $|C\rangle = |\lambda_3, \lambda_3\rangle$ depends on whether these particles are bosons, in which case this state exists and counts as one state, or if they are fermions, in which case this state does not exist and should therefore not be counted in the partition function.

### 3.4.4 The Maxwell-Boltzmann approximation

The above discussion regarding states $|A\rangle$ and $|B\rangle$ points to a simple "fix" of the indistinguishability problem. Since states where two particles are exchanged should be counted only once, we may simply divide the partition function $Z = z^N$ of the distinguishable case by the number of permutations of a set of $N$ particles, i.e., $N!$. This is the **Maxwell-Boltzmann approximation**

$$Z_{\mathrm{MB}} = \frac{z^N}{N!} \tag{3.61}$$

Considering the discussion of state $|C\rangle$, it appears clearly that this will be a good enough approximation if there is little chance that two particles may occupy the same individual state, so if the average number of particles $\langle N_\lambda\rangle$ in a given individual state $|\lambda\rangle$ is small, i.e.,

$$\langle N_\lambda\rangle \ll 1 \tag{3.62}$$

4. We recall that pure quantum states $|\psi\rangle$ are such that the applications of the $\mathcal{S}^2$ and $\mathcal{S}_z$ operators (squared spin modulus and spin projection on a given axis) give $\mathcal{S}^2|\psi\rangle = s(s+1)\hbar^2|\psi\rangle$ and $\mathcal{S}_z|\psi\rangle = s_z\hbar|\psi\rangle$ with $s \geqslant 0$ the spin of the particle, and $s_z$ takes one of $2s+1$ possible values between $-s$ and $+s$.

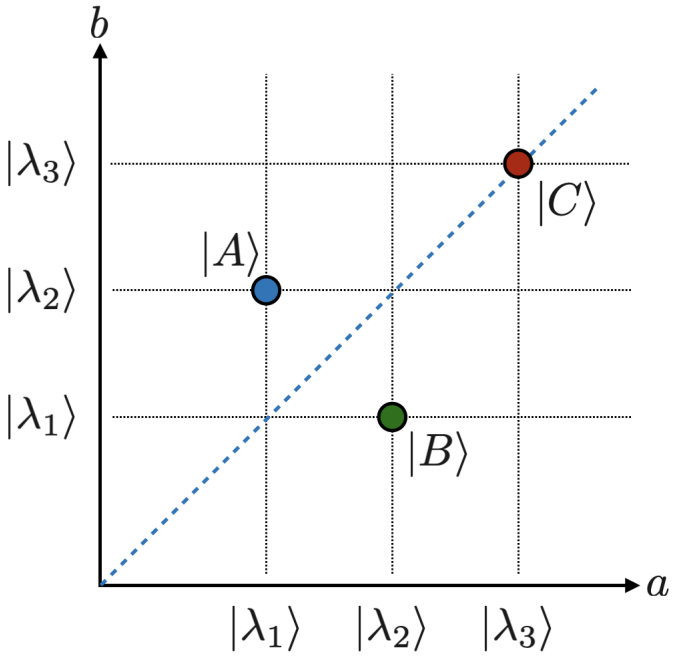



FIGURE 3.4 – Counting states in the indistinguishable case.

Since in a macroscopic system, we have $N \gg 1$ this means that in practice there should be infinitely many individual states for this approximation to be valid. Consequently, it will not apply to systems where the temperature is too low, because in that case the lower-energy levels will be over-populated. So the Maxwell-Boltzmann approximation is valid at high temperature only.

To be more quantitative, we note that the probability to find a given particle in a given individual state $|\lambda\rangle$ is

$$p_\lambda = \frac{e^{-\beta\epsilon_\lambda}}{z} \tag{3.63}$$

so that the average number of particles in that state is

$$\langle N_\lambda \rangle = N p_\lambda = \frac{N e^{-\beta\epsilon_\lambda}}{z} \tag{3.64}$$

For this to be much smaller than one for any individual state, it suffices and it is necessary that it is the case for the fundamental level, whose energy is $\epsilon_0$, i.e.,

$$\frac{N e^{-\beta\epsilon_0}}{z} \ll 1 \tag{3.65}$$

which may be rewritten as follows,

$$\sum_{|\lambda\rangle} e^{-\beta(\epsilon_\lambda - \epsilon_0)} \gg N \tag{3.66}$$

showing that there should be a lot of states (many more terms than $N$ in the sum), and that $\beta$ should be small (for these terms to be non-negligible), i.e., the temperature should be large.

### 3.4.5 Example : the monoatomic ideal gas

.

As an example, we consider an ideal gas of $N$ identical, independent, indistinguishable particles, within a box of volume $V$ in thermal contact with a thermostat at temperature $T$. The Hamiltonian of a single particle in this case is just the kinetic part

$$H_i = \frac{p_i^2}{2m} \tag{3.67}$$

and the individual states are characterized by the momentum or the wave vector, which are quantized,

$$\vec{k}_i = \frac{2\pi n_{i,x}}{L_x}\vec{e}_x + \frac{2\pi n_{i,y}}{L_y}\vec{e}_y + \frac{2\pi n_{i,z}}{L_z}\vec{e}_z \tag{3.68}$$

and the spin projection $s_{i,z}$ which may take one of $2s+1$ values, so we may write symbolically $|\lambda_i\rangle = \left|\vec{k}_i, s_{i,z}\right\rangle$. The energy of this state is

$$\epsilon_{\lambda_i} = \frac{p_i^2}{2m} = \frac{\hbar^2 k_i^2}{2m} \tag{3.69}$$

**Partition function, de Broglie wavelength, and Maxwell-Boltzmann approximation**

The single-particle partition function is then the discrete sum

$$z_i = \sum_{|\lambda_i\rangle} e^{-\beta\epsilon_{\lambda_i}} \tag{3.70}$$

and may be written as an integral over the individual energy $\epsilon$, by introducing the density of states already computed in (1.65), with the additional factor $2s+1$ due to the spin states,

$$\rho(\epsilon) = \frac{(2s+1)Vm^{3/2}}{\sqrt{2}\pi^2\hbar^3}\sqrt{\epsilon} \tag{3.71}$$

We then have the single-particle partition function (assuming $\epsilon_0 = 0$)

$$z = \int_0^\infty \rho(\epsilon) e^{-\beta\epsilon} \mathrm{d}\epsilon \tag{3.72}$$

We leave it to the reader to show [5] that this expression amounts to

$$z = \frac{(2s+1)V}{\Lambda^3} \tag{3.77}$$

where we have introduced the **de Broglie thermal wavelength**

$$\Lambda = \left(\frac{2\pi\hbar^2}{mk_BT}\right)^{1/2} = \frac{h}{\sqrt{2\pi m k_B T}} \tag{3.78}$$

5. The demonstration is as follows, starting from the integral expression

$$z = \int_0^\infty AV\sqrt{\epsilon}e^{-\beta\epsilon}\mathrm{d}\epsilon = AV\int_0^\infty \sqrt{\frac{x}{\beta}}e^{-x}\frac{\mathrm{d}x}{\beta} = \frac{AV}{\beta^{3/2}}\int_0^\infty \sqrt{x}e^{-x}\mathrm{d}x \tag{3.73}$$

where

$$A = \frac{(2s+1)m^{3/2}}{\sqrt{2}\pi^2\hbar^3} \tag{3.74}$$

We then use the integral

$$\int_0^\infty \sqrt{x}e^{-x}\mathrm{d}x = \frac{\sqrt{\pi}}{2} \tag{3.75}$$

to obtain

$$z = AV\left(k_BT\right)^{3/2}\frac{\sqrt{\pi}}{2} = \frac{(2s+1)m^{3/2}}{(2\pi)^{3/2}\hbar^3}V\left(k_BT\right)^{3/2} = \frac{(2s+1)V}{\Lambda^3} \tag{3.76}$$

The Maxwell-Boltzmann approximation then leads to the system's partition function

$$Z = \frac{1}{N!}\frac{(2s+1)^N V^N}{\Lambda^{3N}} \tag{3.79}$$

a result that is valid provided the condition of small average occupation number is met, that is,

$$ze^{\beta\epsilon_0} = z \gg N \tag{3.80}$$

which can be tellingly restated as

$$d = \left(\frac{V}{N}\right)^{1/3} \gg \Lambda \tag{3.81}$$

i.e., the typical separation $d$ between two particles is large compared to the de Broglie thermal wavelength, which characterizes the spatial extent of the particles under random thermal motions [6]. This is also expressible in terms of a condition on the temperature

$$k_B T \gg \frac{2\pi\hbar^2}{m}\left(\frac{N}{V}\right)^{2/3} \tag{3.82}$$

**Properties of the (classical) ideal gas**

From the partition function, we derive the Helmholtz free energy using Stirling's approximation

$$F = -k_B T \ln Z = -Nk_B T\left[\ln\left(\frac{V}{N}\right) + \frac{3}{2}\ln\left(\frac{mk_B T}{2\pi\hbar^2}\right) + 1 + \ln(2s+1)\right] \tag{3.83}$$

and from there the average energy using (3.31), finding the usual result,

$$\langle E\rangle = \frac{3}{2}Nk_B T \tag{3.84}$$

which may be interpreted as there being an average (translational) energy $\langle\epsilon\rangle = (3/2)k_B T$ per particle, i.e., $k_B T/2$ per degree of freedom. This leads to the heat capacity at constant volume

$$C_v = \frac{\partial\langle E\rangle}{\partial T} = \frac{3}{2}Nk_B \tag{3.85}$$

The entropy is now obtained through

$$S = \frac{\langle E\rangle - F}{T} \tag{3.86}$$

and is given by the following **Sackur-Tetrode formula**, which is extensive, showing that the Maxwell-Boltzmann approximation solves the Gibbs paradox,

$$S = Nk_B\left[\ln\left(\frac{V}{N}\right) + \frac{3}{2}\ln\left(\frac{mk_B T}{2\pi\hbar^2}\right) + \frac{5}{2} + \ln(2s+1)\right] \tag{3.87}$$

6. Indeed, this is the $\delta x$ computed from Heisenberg's uncertaintly relation if $\delta p = m\delta v$ is taken to be the typical momentum for thermal motions at temperature $T$.

We recognize that the entropy, mean energy and free energy are linked by the relation known from classical thermodynamics

$$F = \langle E \rangle - TS \tag{3.88}$$

By computing the pressure from the free energy, we recover the usual equation of state

$$P = -\frac{\partial F}{\partial V} = \frac{N k_B T}{V} \tag{3.89}$$

and the chemical potential is given by

$$\mu = \frac{\partial F}{\partial N} = -k_B T \ln\left(\frac{z}{N}\right) \tag{3.90}$$

## 3.5 Canonical distribution in classical mechanics

### 3.5.1 General results

In classical mechanics, a macroscopic system $\mathcal{S}$ of $N$ particles is described by $3N$ **generalized coordinates** $q_i$ and $3N$ **conjugate momenta** $p_i$, forming a $6N$-dimensional phase space. For instance, these coordinates may be Cartesian coordinates $(x_i, y_i, z_i)$, and the conjugate momenta are then $(mv_{i_x}, mv_{i_y}, mv_{i_z})$, but this is not necessarily always the case : in a pendulum, for instance, the most adapted generalized coordinate is the angle with respect to the vertical.

In this context, the equivalent of the Hamiltonian is the classical **Hamilton function** $\mathcal{H}(q_i, p_i)$ and the probability $\mathrm{d}P$ for the system to be in a state in the vicinity of $\{q_i, p_i\}$ is

$$\mathrm{d}P = w\left(\{q_i, p_i\}\right) \mathrm{d}q_1 \ldots \mathrm{d}q_{3N} \mathrm{d}p_1 \ldots \mathrm{d}p_{3N} \tag{3.91}$$

The **phase space density** $w\left(\{q_i, p_i\}\right)$ is the classical equivalent of the probability $p_\ell$ for a quantum state $|\ell\rangle$, and is proportional to $e^{-\beta\mathcal{H}}$ in the canonical approach, so

$$w\left(\{q_i, p_i\}\right) = \frac{e^{-\beta\mathcal{H}}}{A} \tag{3.92}$$

where $A$ is a constant that stems for the normalization condition for the probabilities,

$$A = \int \mathrm{d}q_1 \ldots \mathrm{d}q_{3N} \mathrm{d}p_1 \ldots \mathrm{d}p_{3N} e^{-\beta\mathcal{H}(q_i, p_i)} \tag{3.93}$$

That constant is undoubtedly related to what we would call a classical partition function $Z$, and in fact should be proportional to it. However, it is not exactly the same, since $Z$ is dimensionless, while $A$ has the dimensions of an action to the power $3N$. So there must exist a constant with the proper dimensions that allows to convert from one to the other. We admit that this constant is $h^{3N}$, which may be seen as the elementary volume in phase space. Consequently, in classical statistical mechanics in the canonical approach, we have

$$Z = \frac{1}{(2\pi\hbar)^{3N}} \int \mathrm{d}q_1 \ldots \mathrm{d}q_{3N} \mathrm{d}p_1 \ldots \mathrm{d}p_{3N} e^{-\beta\mathcal{H}(q_i, p_i)} \tag{3.94}$$

Now, this is only valid for a system in which particles are distinguishable. Indeed, taking the case of independent particles, Hamilton's function is separable into

$$\mathcal{H}(q_i, p_i) = \mathcal{H}_1\left(\vec{q}_1, \vec{p}_1\right) + \ldots + \mathcal{H}_N\left(\vec{q}_N, \vec{p}_N\right) \tag{3.95}$$

In that case, the partition function computed above reads

$$Z = \left[ \frac{1}{(2\pi\hbar)^3} \int \mathrm{d}^3\vec{q}\mathrm{d}^3\vec{p} e^{-\beta\mathcal{H}_i(\vec{q},\vec{p})} \right]^N = z^N \tag{3.96}$$

and we would get a non-extensive expression for the entropy, the already mentioned Gibbs paradox. The solving of this paradox in classical statistical mechanics is to divide by $N!$ as was done in the Maxwell-Boltzmann approximation for the quantum treatment. It should be noted that this is not an approximation here, because the probability for two particles to be in the exact same state is rigourously zero since these states are described by continuous variables. Therefore, for a system of $N$ identical and indistinguishable particles, we have the partition function

$$Z = \frac{1}{N!} \frac{1}{(2\pi\hbar)^{3N}} \int \mathrm{d}q_1 \dots \mathrm{d}q_{3N} \mathrm{d}p_1 \dots \mathrm{d}p_{3N} e^{-\beta\mathcal{H}(q_i,p_i)} \tag{3.97}$$

If, moreover, the particles are independent, this factorizes into

$$Z = \frac{z^N}{N!} \tag{3.98}$$

where the single-particle partition function is

$$z = \frac{1}{(2\pi\hbar)^3} \int \mathrm{d}^3\vec{q}\mathrm{d}^3\vec{p} e^{-\beta\mathcal{H}_i(\vec{q},\vec{p})} \tag{3.99}$$

Specializing even further, and considering that we deal with a perfect gas, Hamilton's function is just the kinetic energy, so

$$z = \frac{1}{(2\pi\hbar)^3} \int \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z\, \mathrm{d}p_x\, \mathrm{d}p_y\, \mathrm{d}p_z \exp\left(-\frac{\beta p^2}{2m}\right) = \frac{V}{(2\pi\hbar)^3} \left[ \int \mathrm{d}p_i \exp\left(-\frac{\beta p_i^2}{2m}\right) \right]^3 \tag{3.100}$$

Performing the Gaussian integration, we obtain [7]

$$z = \frac{V}{\Lambda^3} \tag{3.101}$$

where $\Lambda$ is the thermal de broglie wavelength already introduced. This classical treatment is valid if characteristic actions in the problem are much larger than $\hbar$. Here, the typical length is the distance between particles $\tilde{l} \sim (V/N)^{1/3}$ while the typical speed is the thermal one $\tilde{v} \sim \sqrt{3k_BT/m}$ so the typical action is

$$\tilde{l} \times m\tilde{v} \sim \sqrt{mk_BT} \left(\frac{V}{N}\right)^{1/3} \gg \hbar \tag{3.102}$$

so this is the same condition as that for which the Maxwell-Botlzmann approximation is valid.

### 3.5.2 Energy equipartition

Assume that Hamilton's function is quadratic in some momentum $p_m$, that is, it may be written as

$$\mathcal{H}(q_i, p_i) = ap_m^2 + b \tag{3.103}$$

7. Note that we did not consider the particles to have any internal structure, and in particular we ignore their spin, which explains the lack of the factor $2s + 1$.

where $a$ and $b$ might be functions of the other dimensions of phase space $q_i$ and $p_i$, but not of $p_m$. We may show, and the reader is invited to do so as an exercise, that the canonical average of the quadratic term is simply $k_BT/2$, i.e.,

$$\langle ap_m^2 \rangle = \frac{1}{A} \int \prod_{i=1}^{3N} \mathrm{d}q_i \, \mathrm{d}p_i a p_m^2 e^{-\beta \mathcal{H}(q_i,p_i)} = \frac{k_BT}{2} \tag{3.104}$$

With the same reasoning, we find that if Hamilton's function is quadratic in some generalized coordinate $q_m$, i.e. $\mathcal{H}(q_i, p_i) = a'q_m^2 + b'$, where, similarly, $a'$ and $b'$ might be functions of the other dimensions of phase space $q_i$ and $p_i$, but not of $q_m$, then

$$\langle a'q_m^2 \rangle = \frac{1}{A} \int \prod_{i=1}^{3N} \mathrm{d}q_i \, \mathrm{d}p_i a' q_m^2 e^{-\beta \mathcal{H}(q_i,p_i)} = \frac{k_BT}{2} \tag{3.105}$$

This is what is meant by saying usually that each degree of freedom carries a typical thermal energy $k_BT/2$. It may be applied to an ideal monoatomic gas, since its energy is quadratic in all momenta, separately, and there are $3N$ such momenta for a system of $N$ particles, so that

$$\langle E \rangle = \langle \mathcal{H} \rangle = \frac{3}{2} N k_B T \tag{3.106}$$

When internal degrees of freedom are thermally accessible, these provide supplementary $k_BT/2$ terms. For instance, in a diatomic ideal gas where the two rotational degrees of freedom are excited,

$$\langle E \rangle = \langle \mathcal{H} \rangle = \frac{5}{2} N k_B T \tag{3.107}$$

As a further example, we consider the case of the one-dimensional harmonic oscillator, for which Hamilton's function is

$$\mathcal{H} = \frac{p_x^2}{2m} + \frac{1}{2} k x^2 \tag{3.108}$$

and therefore the mean energy per oscillator is $\langle \epsilon \rangle = k_BT$. In a system with $3N$ oscillators, such as a crystal containing $N$ atoms, we thus have

$$\langle E \rangle = 3 N k_B T \tag{3.109}$$

and therefore the heat capacity at constant volume in this model is a constant, $C_v = 3Nk_B$, which yields the **Dulong-Petit law**. We note that this is only valid when the potential energy is indeed harmonic, close to the miminum. When moving away from the minimum, anharmonicity effects may render Hamilton's function non-quadractic in the generalized coordinates, and the equipartition theorem no longer applies.

# 4 Grand-canonical ensemble

In this chapter, we discuss the **grand-canonical distribution**, which gives the probabilities of the various microstates of a system that is in contact with a reservoir of both heat and particles, that is a much larger system imposing its temperature and chemical potential to the system. This means that the system is free to exchange energy and particles with the reservoir, but other parameters, in particular the volume, are fixed.

## 4.1 Grand-canonical distribution

### 4.1.1 Reservoir

The system $\mathcal{S}$ under study is assumed to be in contact with a much larger system $\mathcal{R}$ with which it is possible to exchange both energy and particles. We assume that only one type of particles is exchanged between $\mathcal{S}$ and $\mathcal{R}$, and that the complete system $\mathcal{S} \cup \mathcal{R}$ is isolated. As a consequence, the total energy and number of particles are fixed quantities

$$E + E_{\mathcal{R}} = E_0 \qquad N + N_{\mathcal{R}} = N_0 \tag{4.1}$$

The system $\mathcal{R}$ is a particle and energy reservoir if $E_{\mathcal{R}} \gg E$ and $N_{\mathcal{R}} \gg N$ for all values of $E$ and $N$ that have non-negligible probabilities of occurring.

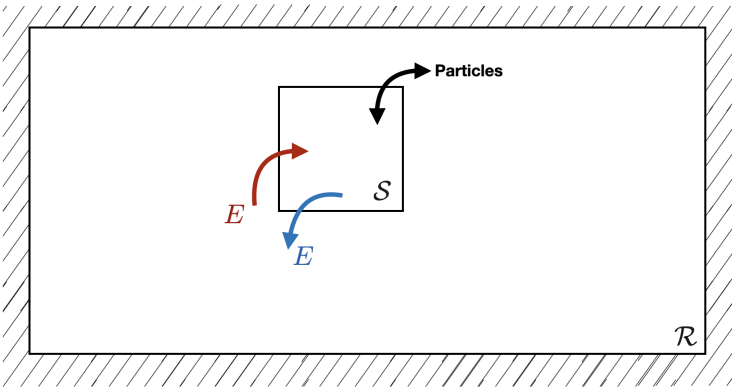



FIGURE 4.1 – Grand-canonical setup

**Microcanonical entropy of the reservoir**

The microcanonical entropy of the reservoir may be computed to first order in $E$ and $N$,

$$S_{\mathcal{R}}(E_0 - E, N_0 - N) = S_{\mathcal{R}}(E_0, N_0) - E\frac{\partial S_{\mathcal{R}}}{\partial E_{\mathcal{R}}}(E_0, N_0) - N\frac{\partial S_{\mathcal{R}}}{\partial N_{\mathcal{R}}}(E_0, N_0) \tag{4.2}$$

which allows to introduce the microcanonical temperature $T_{\mathcal{R}}$ and chemical potential $\mu_{\mathcal{R}}$ of the reservoir, both taken at the values $(E_0, N_0)$ of the reservoir's energy and particle number

$$S_{\mathcal{R}}(E_0 - E, N_0 - N) = S_{\mathcal{R}}(E_0, N_0) - \frac{E}{T_{\mathcal{R}}} + \frac{\mu_{\mathcal{R}}}{T_{\mathcal{R}}}N \tag{4.3}$$

The system $\mathcal{R}$ acts as a proper reservoir if $\mu_{\mathcal{R}}$ and $T_{\mathcal{R}}$ do not depend on the values of the energy $E$ and number of particles $N$ in the system $\mathcal{S}$. This means that, for instance,

$$T_{\mathcal{R}}(E_0 - E, N_0 - N) = T_{\mathcal{R}}(E_0, N_0) - E\frac{\partial T_{\mathcal{R}}}{\partial E_{\mathcal{R}}}(E_0, N_0) - N\frac{\partial T_{\mathcal{R}}}{\partial N_{\mathcal{R}}}(E_0, N_0) \approx T_{\mathcal{R}}(E_0, N_0) \tag{4.4}$$

Now the terms to be neglected are indeed much smaller than the zeroth-order term, provided $N \ll N_{\mathcal{R}}$,

$$E\frac{\partial T_{\mathcal{R}}}{\partial E_{\mathcal{R}}}(E_0, N_0) \sim \frac{N}{N_{\mathcal{R}}}T_{\mathcal{R}} \ll T_{\mathcal{R}} \qquad N\frac{\partial T_{\mathcal{R}}}{\partial N_{\mathcal{R}}}(E_0, N_0) \sim \frac{N}{N_{\mathcal{R}}}T_{\mathcal{R}} \ll T_{\mathcal{R}} \tag{4.5}$$

A perfect reservoir, in the limit $N_{\mathcal{R}} \to \infty$ would have constant temperature and chemical potential. We may regard these as fixed external parameters and write $T = T_{\mathcal{R}}$ and $\mu = \mu_{\mathcal{R}}$ to stand for the **grand-canonical temperature and chemical potential**, imposed by the reservoir to the system $\mathcal{S}$.

**The different statistical ensembles**

We recall that in the microcanonical approach, $E$, $V$, and $N$ were fixed, and the fundamental quantity was the microcanonical entropy $S(E, V, N)$, from which we derived (as we just did again), the microcanonical temperature and chemical potential. In the canonical approach, the fixed parameters were $T$, $V$, and $N$, and the fundamental quantity was the Helmholtz free energy $F(T, V, N)$ from which we derived quantities such as the average energy $\langle E \rangle$. Now, in the grand-canonical approach, the fixed parameters are $T$, $V$, and $\mu$, and it is reasonable to ask what the fundamental quantity is from which me may derive an average energy $\langle E \rangle$ or number of particles $\langle N \rangle$.

### 4.1.2 Grand-canonical distribution

The system $\mathcal{S}$ is in equilibrium with $\mathcal{R}$, that imposes its temperature and chemical potential. We look for the probabilities $p_\ell$ that the system is in a given microstate $|\ell\rangle$, which is characterized in particular by an energy $E_\ell$ and a number of particles $N_\ell$.

We assume that the coupling between $\mathcal{S}$ and $\mathcal{R}$ is weak. A microstate of $\mathcal{S} \cup \mathcal{R}$ is written $|\ell, L\rangle$ where $|\ell\rangle$ is the microstate of $\mathcal{S}$ and $|L\rangle$ is that of $\mathcal{R}$. The probability of that combined microstate is given by the microcanonical distribution, since $\mathcal{S} \cup \mathcal{R}$ is isolated,

$$p_{\ell,L} = \frac{1}{\Omega_{\mathcal{S}\cup\mathcal{R}}(E_0, N_0)} \tag{4.6}$$

but what we are after is the probability $p_\ell$, and in this state of $\mathcal{S}$ the energy of the reservoir must be $E_L = E_0 - E_\ell$, within the experimental uncertainty $\delta E$ on the total energy, and the number of particles in the reservoir must be $N_L = N_0 - N_\ell$. There is no "uncertainty" in this latter case because numbers of particles are discrete quantities. Consequently,

$$p_\ell = \frac{\Omega_{\mathcal{R}}(E_0 - E_\ell, N_0 - N_\ell)}{\Omega_{\mathcal{S}\cup\mathcal{R}}(E_0, N_0)} \propto \exp\left[\frac{S_{\mathcal{R}}(E_0 - E_\ell, N_0 - N_\ell)}{k_B}\right] \tag{4.7}$$

The Taylor expansion of the microcanonical entropy of $\mathcal{R}$ was already computed in (4.3) so we have

$$p_\ell \propto \exp\left(-\frac{E_\ell}{k_B T} + \frac{\mu N_\ell}{k_B T}\right) \tag{4.8}$$

Using the normalization condition, we finally write the **grand-canonical distribution** as

$$p_\ell = \frac{1}{\Xi} \exp\left(-\frac{E_\ell}{k_B T} + \frac{\mu N_\ell}{k_B T}\right) \tag{4.9}$$

where we have introduced the **grand-canonical partition function**

$$\Xi = \sum_{|\ell\rangle} \exp\left(-\frac{E_\ell}{k_B T} + \frac{\mu N_\ell}{k_B T}\right) \tag{4.10}$$

### 4.1.3 Internal variables

#### Number of particles

Let us consider the probability $p(N)$ that the system's number of particles is $N_\ell = N$. This may be computed from the sum of $p_\ell$, restricted to the appropriate set of states,

$$p(N) = \sum_{\{|\ell\rangle | N_\ell = N\}} p_\ell = \frac{1}{\Xi} \sum_{\{|\ell\rangle | N_\ell = N\}} e^{-\beta(E_\ell - \mu N_\ell)} = \frac{e^{\beta\mu N}}{\Xi} \sum_{\{|\ell\rangle | N_\ell = N\}} e^{-\beta E_\ell} \tag{4.11}$$

The remaining sum is obviously the canonical partition function for the chosen number of particles $N$,

$$p(N) = \frac{e^{\beta\mu N}}{\Xi} Z(T, V, N) \tag{4.12}$$

Now we may ask what is the most probable value of $N$. This value $N_m$ is the one maximizing $p(N)$ found above, so it is also the one that maximizes $\ln[p(N)]$. Since $\Xi$ does not depend on $N$, this is equivalent to finding the value maximizing

$$\beta\mu N + \ln Z = \beta(\mu N - F) \tag{4.13}$$

since $F = -k_B T \ln Z$. The equation to solve is then

$$\frac{\partial}{\partial N}(\mu N - F) = \mu - \frac{\partial F}{\partial N} = 0 \tag{4.14}$$

which simply means that $N_m$ is the value of $N$ for which the canonical chemical potential $\mu_c(T, V, N_m)$, defined in (3.39), is equal to the chemical potential $\mu$ imposed by the reservoir. This may be (at least formally) inverted to give $N_m$ as a function of $T$, $V$, and $\mu$.

As usual, if the system $\mathcal{S}$ is macroscopic, the distribution of $p(N)$ is Gaussian around the mean and most probable value $N_m$, with

$$p(N) = \frac{1}{\sqrt{2\pi}\sigma_N} \exp\left[-\frac{(N - N_m)^2}{2\sigma_N^2}\right] \tag{4.15}$$

where the variance is given by the following formula, which may be worked out as an exercise,

$$\sigma_N^2 = k_B T \left[ \left( \frac{\partial^2 F}{\partial N^2} \right)_{T,V} (N_m) \right]^{-1} \tag{4.16}$$

The relative fluctuations of the number of particles in the system are of the order

$$\frac{\sigma_N}{N_m} \sim \frac{1}{\sqrt{N_m}} \ll 1 \tag{4.17}$$

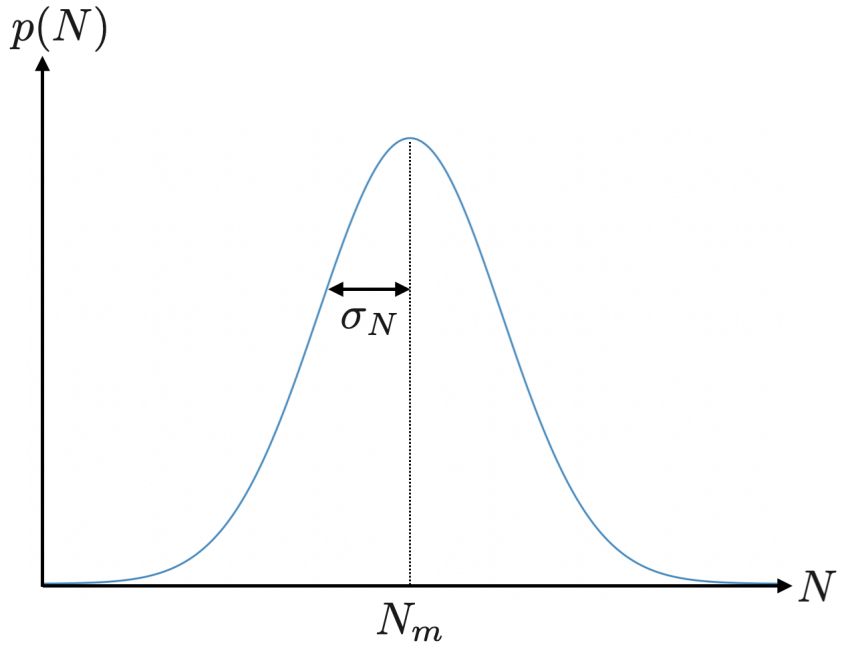



FIGURE 4.2 – Distribution of the total number of particles in the grand-canonical approach

**Other internal variables**

For any other internal variable $Y$, the probability that it should take a given value $y$ is given by

$$p(y) = \frac{1}{\Xi} \sum_{\{|\ell\rangle | y_\ell = y\}} e^{-\beta(E_\ell - \mu N_\ell)} \tag{4.18}$$

but this does not lend itself to any simpler form in the general case. It must be computed specifically for each variable.

## 4.2 Grand-canonical partition function and potential

### 4.2.1 Definitions

As already stated, the **grand-canonical partition function** is defined by

$$\Xi(T, V, \mu) = \sum_{|\ell\rangle} \exp\left( -\frac{E_\ell}{k_B T} + \frac{\mu N_\ell}{k_B T} \right) \tag{4.19}$$

This can be re-written in terms of the canonical partition function $Z(T,V,N)$ as a sum over $N$,

$$\Xi(T,V,\mu) = \sum_N \sum_{\{|\ell\rangle | N_\ell = N\}} e^{-\beta E_\ell} e^{\beta \mu N} = \sum_N \varphi^N Z(T,V,N) \tag{4.20}$$

where we have introduced the **fugacity** $\varphi = e^{\beta\mu}$. The sum in the above equation only has a theoretical interest, since if we know the canonical partition function $Z(T,V,N)$ we already are able to derive all there is to know about the system.

From the grand-canonical partition function, we may define the **grand-canonical potential** as[1]

$$J(T,V,\mu) = -k_B T \ln \Xi \tag{4.21}$$

### 4.2.2 Properties of the system

Using the quantities defined above, we may derive the average number of particles and the average energy of the system, their standard deviations around the mean, as well as define the grand-canonical entropy and pressure.

#### Average number of particles

This is defined as

$$\langle N \rangle = \sum_{|\ell\rangle} N_\ell p_\ell = \frac{1}{\Xi} \sum_{|\ell\rangle} N_\ell e^{-\beta(E_\ell - \mu N_\ell)} = \frac{1}{\Xi} \frac{1}{\beta} \frac{\partial}{\partial \mu} \left[ \sum_{|\ell\rangle} e^{-\beta(E_\ell - \mu N_\ell)} \right] \tag{4.22}$$

which shows that it may be obtained through

$$\langle N \rangle = \frac{1}{\beta} \frac{\partial \ln \Xi}{\partial \mu} = -\frac{\partial J}{\partial \mu} \tag{4.23}$$

#### Average energy

Using the same line of reasoning, we have

$$\frac{\partial \ln \Xi}{\partial \beta} = \langle \mu N - E \rangle = \mu \langle N \rangle - \langle E \rangle \tag{4.24}$$

and therefore the average energy is

$$\langle E \rangle = \mu \langle N \rangle - \frac{\partial \ln \Xi}{\partial \beta} = \frac{1}{\beta} \left[ \mu \frac{\partial}{\partial \mu} - \beta \frac{\partial}{\partial \beta} \right] \ln \Xi \tag{4.25}$$

#### Standard deviations

The standard deviations of the number of particles and the energy are given by

$$\sigma_N^2 = \frac{1}{\beta^2} \frac{\partial^2 \ln \Xi}{\partial \mu^2} \qquad \sigma_E^2 = \left[ \frac{\mu}{\beta} \frac{\partial}{\partial \mu} - \frac{\partial}{\partial \beta} \right]^2 \ln \Xi \tag{4.26}$$

The reader is invited to obtain these expressions as an exercise.

1. Note that in some texts, it is written $\Psi(T,V,\mu)$ or $\Omega(T,V,\mu)$ (not to be confused with the number of accessible states in the microcanonical approach).

**Grand-canonical entropy**

To determine the expression of the grand-canonical entropy, we start from the general expression for the entropy, in which we specify the values of the grand-canonical probabilities

$$S = -k_B \sum_{|\ell\rangle} p_\ell \ln p_\ell = -k_B \sum_{|\ell\rangle} \frac{1}{\Xi} e^{-\beta(E_\ell - \mu N_\ell)} \left[\beta\left(\mu N_\ell - E_\ell\right) - \ln \Xi\right] \tag{4.27}$$

From this, we recognize the average number of particles, average energy, and we have

$$S = \frac{1}{T}\left(\langle E\rangle - \mu\langle N\rangle\right) + k_B \ln \Xi = \frac{\langle E\rangle - \mu\langle N\rangle - J}{T} \tag{4.28}$$

This may be written in a form that is more familiar from classical thermodynamics

$$J = \langle E\rangle - TS - \mu\langle N\rangle \tag{4.29}$$

**Infinitesimal variation of $J$**

We have the following partial derivatives (that may be computed as an exercise)

$$\left(\frac{\partial J}{\partial T}\right)_{V,\mu} = -S \qquad \left(\frac{\partial J}{\partial \mu}\right)_{V,T} = -\langle N\rangle \tag{4.30}$$

and we define the **grand canonical pressure** as

$$P = -\left(\frac{\partial J}{\partial V}\right)_{T,\mu} \tag{4.31}$$

This pressure is equal to the microcanonical and canonical pressures in macroscopic systems. This allows to write the infinitesimal variation of $J$ as a function of those of its natural variables $T$, $V$, and $\mu$,

$$\mathrm{d}J = -S\mathrm{d}T - P\mathrm{d}V - \langle N\rangle \mathrm{d}\mu \tag{4.32}$$

### 4.2.3 The case of two separate or weakly coupled sub-systems

Consider a system $\mathcal{S}$ consisting of two sub-systems $\mathcal{S}_1$ and $\mathcal{S}_2$, that are separated by a fixed wall, and weakly coupled. The whole system is in contact with a reservoir $\mathcal{R}$, that imposes its temperature and chemical potential, so we have $\beta_1 = \beta_2 = \beta$ and $\mu_1 = \mu_2 = \mu$. The state $|\ell\rangle$ of the system $\mathcal{S}$ may be thought of as a pair of states for each of the two sub-systems, i.e., $|\ell\rangle = |\ell_1, \ell_2\rangle = |\ell_1\rangle \otimes |\ell_2\rangle$. The energy and number of particles of the system are simply

$$E_\ell = E_{\ell_1} + E_{\ell_2} \qquad N_\ell = N_{\ell_1} + N_{\ell_2} \tag{4.33}$$

and the grand canonical partition function then reads

$$\Xi = \sum_{|\ell\rangle} e^{-\beta(E_\ell - \mu N_\ell)} = \sum_{|\ell_1\rangle} \sum_{|\ell_2\rangle} e^{-\beta(E_{\ell_1} - \mu N_{\ell_1})} e^{-\beta(E_{\ell_2} - \mu N_{\ell_2})} = \Xi_1 \Xi_2 \tag{4.34}$$

which leads to the additivity of the grand-canonical potential

$$J = J_1 + J_2 \tag{4.35}$$

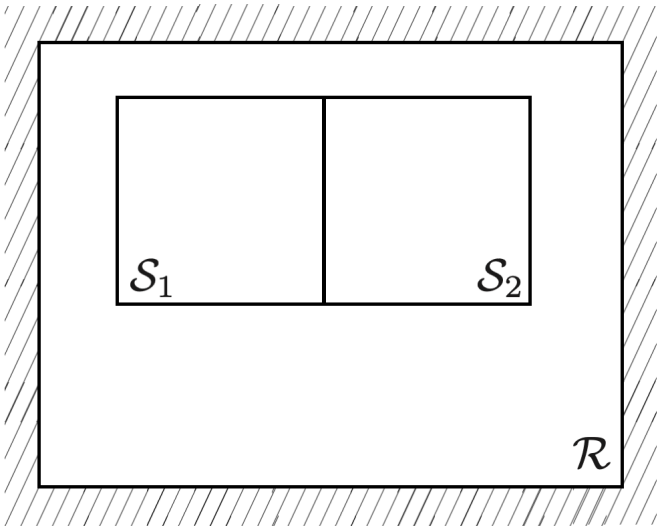



FIGURE 4.3 – Weakly coupled sub-systems in the grand-canonical framework

As an application, we may consider the case of a simple fluid, whose thermodynamic state is completely described by $T$, $V$, and $\mu$, and show that its grand-canonical potential has a very simple expression. Consider two equal volumes $V$ within the system $\mathcal{S}$. The additivity just demonstrated implies that $J(T, 2V, \mu) = 2J(T, V, \mu)$. This means that, generally, for any $\alpha \geqslant 0$, we have $J(T, \alpha V, \mu) = \alpha J(T, V, \mu)$. In other words, $J$ seen as a function of $V$ (with $T$ and $\mu$ fixed) is a homogeneous function of order one. This relation may be derived with respect to $\alpha$,

$$\frac{\partial}{\partial \alpha} J(T, \alpha V, \mu) = V \times \frac{\partial J}{\partial V}(T, \alpha V, \mu) = -P(T, \alpha V, \mu)V = \frac{\partial}{\partial \alpha}\left[\alpha J(T, V, \mu)\right] = J(T, V, \mu) \tag{4.36}$$

which gives the following simple expression for the grand-canonical potential

$$J = -PV \tag{4.37}$$

and also shows that the pressure does not depend on $V$. In that case, the average energy takes the form

$$\langle E \rangle = TS - PV + \mu \langle N \rangle \tag{4.38}$$

### 4.2.4 Relaxation of a constraint

We return to the situation of a system $\mathcal{S}$ in equilibrium with $\mathcal{R}$, that imposes its temperature and chemical potential, and assume that some constraint on $\mathcal{S}$ is relaxed. The entropy of the full isolated system $\mathcal{S} \cup \mathcal{R}$ will necessarily increase. We ask what the evolution of the grand-canonical potential $J$ of the system $\mathcal{S}$ is. To answer this, we write out the microcanonical entropy of the complete system

$$S_{\mathcal{S}\cup\mathcal{R}} = -k_B \sum_{|\ell, L\rangle} p_{\ell,L} \ln p_{\ell,L} \tag{4.39}$$

where the sum extends over all microstates $|\ell, L\rangle$ of the complete system and these states' probabilities are $p_{\ell,L} = p_\ell p_L(E_0 - E_\ell, N_0 - N_\ell)$, with $E_0 = E_\ell + E_L$ the total energy of the system and $N_0 = N_\ell + N_L$

the total number of particles. From this we have

$$S_{\mathcal{S}\cup\mathcal{R}} = -k_B \sum_{|\ell,L\rangle} p_\ell p_L(E_0 - E_\ell, N_0 - N_\ell) \left[\ln p_\ell + \ln \left[p_L(E_0 - E_\ell, N_0 - N_\ell)\right]\right] \tag{4.40}$$

which we may rewrite by separating the sums over the states of $\mathcal{S}$ and the states of $\mathcal{R}$ (we omit the dependence of $p_L$ on $E_0 - E_\ell$ and $N_0 - N_\ell$ for readability)

$$S_{\mathcal{S}\cup\mathcal{R}} = -k_B \sum_{|\ell\rangle} p_\ell \ln p_\ell \sum_{|L\rangle} p_L - k_B \sum_{|\ell\rangle} p_\ell \sum_{|L\rangle} p_L \ln p_L \tag{4.41}$$

The first sum in the first term is related to the entropy of the system $\mathcal{S}$, the second sum is unity due to the normalization condition for the states of the reservoir compatible with the energy and number of particles of $\mathcal{S}$, and the second sum in the second term is related to the entropy of the reservoir $\mathcal{R}$, so

$$S_{\mathcal{S}\cup\mathcal{R}} = S + \sum_{|\ell\rangle} p_\ell S_{\mathcal{R}}(E_0 - E_\ell, N_0 - N_\ell) \tag{4.42}$$

In that last sum, the entropy $S_{\mathcal{R}}$ may be developed in the vicinity of $(E_0, N_0)$ since $E_\ell \ll E_0$ and $N_\ell \ll N_0$, so

$$S_{\mathcal{S}\cup\mathcal{R}} = S + \sum_{|\ell\rangle} p_\ell \left[S_{\mathcal{R}}(E_0, N_0) - \frac{E_\ell}{T} + \frac{\mu N_\ell}{T}\right] = S + S_{\mathcal{R}}(E_0, N_0) - \frac{\langle E\rangle}{T} + \frac{\mu\langle N\rangle}{T} = S_{\mathcal{R}}(E_0, N_0) - \frac{J}{T} \tag{4.43}$$

Since $S_{\mathcal{R}}(E_0, N_0)$ is a constant, the relaxation of an internal constraint leads to an increase of the entropy $S_{\mathcal{S}\cup\mathcal{R}}$ of the complete system, or, equivalently, to a **decrease of the grand-canonical potential** $J$ of the system $\mathcal{S}$ on the path to equilibrium.

$$\mathrm{d}J \leqslant 0 \tag{4.44}$$

## 4.3 The thermodynamic limit

We already saw that for macroscopic systems, there is an equivalence between the microcanonical and canonical descriptions called the thermodynamic limit. By macroscopic systems, we consider that both $V \to \infty$ and $N \to \infty$ while the density $N/V$ remains constant. We recall that the reason for this equivalence lies in the fact that in this limit, the fluctuations of an internal variable become negligibly small, and the Gaussian distribution of its possible values is extremely peaked around the mean (also most probable) value. Therefore, there is virtually nothing to distinguish this variable from an external parameter, whose value is fixed by definition.

In fact, in this limit, all statistical descriptions (microcanonical, canonical, grand-canonical, ...) are equivalent when studying a system at equilibrium, and we may choose the most interesting approach to perform this study.

- In the **microcanonical** approach, $E$, $V$, and $N$ are fixed, the thermodynamic potential of interest is the entropy $S(E, V, N)$, and the microcanonical temperature, pressure, and chemical potentials are internal variables $T^*(E, V, N)$, $P^*(E, V, N)$, and $\mu^*(E, V, N)$.
- In the **canonical** approach, the fixed parameters are $T$, $V$, and $N$, the thermodynamic potential of interest is the Helmholtz free energy $F(T, V, N)$, and we computed the mean energy $\langle E\rangle(T, V, N)$ from equation (3.27) and the canonical pressure $P_c(T, V, N)$ and the canonical chemical potential $\mu_c(T, V, N)$ from Eq. (3.39).

— In the **grand-canonical** approach, the fixed parameters are $T$, $V$, and $\mu$, the thermodynamic potential of interest is the grand-canonical potential $J(T,V,\mu)$, and we computed the mean energy $\langle E\rangle(T,V,\mu)$, mean particle number $\langle N\rangle(T,V,\mu)$, and grand-canonical pressure $P_{\mathrm{GC}}(T,V,\mu)$ from equations (4.25), (4.23), and (4.31), respectively

For a given system, say a monoatomic perfect gas, we can work out its properties such as the equation of state or the relationship between energy, volume, temperature, and pressure in a fashion that depends *a priori* on the experimental conditions. For instance, assuming that the system is isolated we would work in the microcanonical framework and we would (and in fact did) find that

$$T^* = \frac{2E}{3Nk_B} \tag{4.45}$$

Now, we could think of the system as being in contact with a thermostat and compute the mean energy in the canonical framework, finding that

$$\langle E\rangle = \frac{3}{2}Nk_BT \tag{4.46}$$

We see that the equation relating $T^*$, $E$, $V$, and $N$ in the first approach is identical to the one relating $T$, $\langle E\rangle$, $V$, and $N$ in the second approach. Generalizing, we can formally think of the external parameters and internal variables, grouped under $\vec{X}$, as being related in the different statistical descriptions through functions ($\mathcal{F}_{\mathrm{MC}}$, $\mathcal{F}_{\mathrm{C}}$, $\mathcal{F}_{\mathrm{GC}}$) such that

$$\mathcal{F}_{\mathrm{MC}}\left(\vec{X}\right) = 0 \qquad \mathcal{F}_{\mathrm{C}}\left(\vec{X}\right) = 0 \qquad \mathcal{F}_{\mathrm{GC}}\left(\vec{X}\right) = 0 \tag{4.47}$$

The thermodynamic limit states that for macroscopic systems, the fluctuations of internal variables become negligible, implying that these functions become identical

$$\mathcal{F}_{\mathrm{MC}} = \mathcal{F}_{\mathrm{C}} = \mathcal{F}_{\mathrm{GC}} \tag{4.48}$$

## 4.4 Identical, independent, indistinguishable particles

### 4.4.1 General results

#### Setup

We consider a system of independent particles, which means that the Hamiltonian is just a sum of terms, each related to one particle only

$$\mathcal{H} = \sum_{i=1}^{N} H_i \tag{4.49}$$

and we assume furthermore that the particles are identical, which means that all terms of the Hamiltonian are identical $\forall(i,j)\, H_i = H_j$. Given a single particle whose Hamiltonian is $H$, we write $|\lambda\rangle$ for its individual states, and $\epsilon_\lambda$ for their respective energies. For the complete system $\mathcal{S}$, we write $|\ell\rangle$ for the microstates, that are characterized by the numbers $N_\lambda$ of individual particles that are in a given state $|\lambda\rangle$, i.e., $|\ell\rangle = \{N_1, N_2, \ldots, N_\lambda, \ldots\}$. Obviously, we have

$$N_\ell = \sum_{|\lambda\rangle} N_\lambda \qquad\qquad E_\ell = \sum_{|\lambda\rangle} N_\lambda \epsilon_\lambda \tag{4.50}$$

Note that we are already implying that the particles are indistinguishable by stating that a microstate of the system is given by the number of particles in each individual state rather than by the different states of individual particles.

**Grand-canonical partition function**

The expression of the grand-canonical partition function is straightforward to derive

$$\Xi = \sum_{|\ell\rangle} e^{-\beta(E_\ell - \mu N_\ell)} = \sum_{N_1}\sum_{N_2}\dots\sum_{N_\lambda}\dots\left\{\exp\left[-\beta\left(\sum_{|\lambda\rangle} N_\lambda\epsilon_\lambda - \mu\sum_{|\lambda\rangle} N_\lambda\right)\right]\right\} \tag{4.51}$$

The important thing to notice here is that in the grand-canonical framework, we are not constrained by the total number of particles being fixed. In (3.4), we were stopped by this : once we had chosen $N_1$, the number of particles in the ground state, $N_2$ was restricted and depended on the value of $N_1$, which rendered the computation impossible. Here, we may (a priori) choose the number of particles in a given individual state completely independently from the other states. Consequently, the above relation is directly factored into

$$\Xi = \prod_{|\lambda\rangle} \xi_\lambda \tag{4.52}$$

with $\xi_\lambda$ the grand-canonical partition function for the individual state $|\lambda\rangle$

$$\xi_\lambda = \sum_{N_\lambda} e^{-\beta N_\lambda(\epsilon_\lambda - \mu)} \tag{4.53}$$

Note that in the canonical framework, we had found a relation for distinguishable particles, where the factorization was done on the individual particles,

$$Z = \prod_i z_i \tag{4.54}$$

where $z_i$ was a sum over the individual states. Here, the factorization is done on the individual states, and the factors $\xi_\lambda$ are sums over the number of particles occupying that state, and as we mentioned already, the particles are indistinguishable.

**Properties of the system**

From this we may compute the grand-canonical potential

$$J = -k_B T \ln \Xi = -k_B T \sum_{|\lambda\rangle} \ln \xi_\lambda \tag{4.55}$$

the average number of particles in the system, and the average energy

$$\langle N\rangle = \sum_{|\lambda\rangle}\langle N_\lambda\rangle = k_B T \sum_{|\lambda\rangle} \frac{\partial \ln \xi_\lambda}{\partial \mu} \qquad \langle E\rangle = \sum_{|\lambda\rangle}\langle N_\lambda\rangle\epsilon_\lambda \tag{4.56}$$

In these relations, we have exhibited an important quantity that is the **mean occupation number** of an individual state, a quantity that will prove central to quantum gas statistics,

$$\langle N_\lambda\rangle = \frac{1}{\beta}\frac{\partial \ln \xi_\lambda}{\partial \mu} \tag{4.57}$$

### 4.4.2 Quantum statistics

#### Fermions

As already mentioned, fermions are particles with half-integer spin. They obey the Pauli exclusion principle stating that there cannot be more than one fermion to occupy a given individual state. Consequently, for fermions we have only two possibilities

$$N_\lambda = 0 \qquad\qquad N_\lambda = 1 \tag{4.58}$$

The grand-canonical partition function for the individual states is therefore

$$\xi_\lambda^F = 1 + e^{-\beta(\epsilon_\lambda - \mu)} \tag{4.59}$$

and from there the mean occupation number is the **Fermi-Dirac distribution**

$$\langle N_\lambda \rangle^F = \frac{1}{e^{\beta(\epsilon_\lambda - \mu)} + 1} \tag{4.60}$$

We note that, as expected, $\langle N_\lambda \rangle^F < 1$.

#### Bosons

Bosons are particles with integer spin, and there can be any number of bosons occupying a given individual state. Consequently, the grand-canonical partition function for an individual state is

$$\xi_\lambda^B = \sum_{N_\lambda=0}^{\infty} e^{-N_\lambda \beta(\epsilon_\lambda - \mu)} = \sum_{N_\lambda=0}^{\infty} \left[ e^{-\beta(\epsilon_\lambda - \mu)} \right]^{N_\lambda} \tag{4.61}$$

This is an infinite sum that only converges if $e^{-\beta(\epsilon_\lambda - \mu)} < 1$, and this must be true for all individual states $|\lambda\rangle$, so, since $\beta > 0$, we must have $\epsilon_\lambda - \mu > 0$ for all states. This is equivalent to stating that the ground state has an energy larger than the chemical potential [2]

$$\epsilon_0 > \mu \tag{4.62}$$

With that condition, the computation of the infinite sum is straightforward,

$$\xi_\lambda^B = \frac{1}{1 - e^{-\beta(\epsilon_\lambda - \mu)}} \tag{4.63}$$

and from there the mean occupation number is the **Bose-Einstein distribution**

$$\langle N_\lambda \rangle^B = \frac{1}{e^{\beta(\epsilon_\lambda - \mu)} - 1} \tag{4.64}$$

We note that, as expected, $\langle N_\lambda \rangle^B$ may be larger than one in this case.

The important thing to note here is that relaxing the constraint on the total number of particles by working in the grand-canonical framework has allowed us to perform a computation that was altogether impossible in the canonical framework.

2. This is because, should the reverse be true, there would be a constant flux of (ground state energy) particles from the reservoir to the system, effectively preventing any equilibrium. This is connected to the phenomenon of Bose condensation discussed in Chapter 6.

**The Maxwell-Boltzmann approximation**

We recall that the Maxwell-Boltzmann approximation was meant to fix the counting problem of the indistinguishable case in the canonical approach by dividing the partition function $Z$ by $N!$, which is correct provided that the number of particles in a given individual state is $\langle N_\lambda \rangle \ll 1$. Considering the Fermi-Dirac and Bose-Einstein distributions, this is equivalent to having

$$e^{\beta(\epsilon_\lambda - \mu)} \gg 1 \tag{4.65}$$

and in that case the two distributions are very well approximated by the Boltzmann distribution

$$\langle N_\lambda \rangle^F \approx \langle N_\lambda \rangle^B \approx \langle N_\lambda \rangle^{MB} = e^{-\beta(\epsilon_\lambda - \mu)} \tag{4.66}$$

We may show how this approximation is related to the division of the canonical partition function by a factor $N!$. Indeed, the grand-canonical partition functions for the individual states are, in the two cases discussed above,

$$\xi_\lambda^F = \frac{1}{1 - \langle N_\lambda \rangle^F} \qquad \xi_\lambda^B = 1 + \langle N_\lambda \rangle^B \tag{4.67}$$

so that for the full grand-canonical partition function we have

$$\ln \Xi = \sum_{|\lambda\rangle} \ln \xi_\lambda \approx \sum_{|\lambda\rangle} \langle N_\lambda \rangle \approx e^{\beta\mu} \sum_{|\lambda\rangle} e^{-\beta\epsilon_\lambda} = e^{\beta\mu} z \tag{4.68}$$

introducing the canonical partition function $z$ for a single particle. We therefore have

$$\Xi = \exp\left[e^{\beta\mu} z\right] = \sum_{N=0}^{\infty} \frac{z^N}{N!} e^{N\beta\mu} = \sum_{N=0}^{\infty} \frac{z^N}{N!} \varphi^N \tag{4.69}$$

that ought to be compared to equation (4.20), from which we have the expected result

$$Z(T, V, N) = \frac{z^N}{N!} \tag{4.70}$$

### 4.4.3 Grand-canonical description of a classical perfect gas

By "classical", we mean of course that the quantum properties are not relevant, which is simply to state that the Maxwell-Boltzmann approximation holds. In that case, we have

$$\Xi = \exp\left[e^{\beta\mu} z\right] \tag{4.71}$$

with $z$ the canonical partition function for a single particle, that we recall here

$$z = \frac{(2s+1)V}{\Lambda^3} \tag{4.72}$$

where we use the de Broglie thermal wavelength

$$\Lambda = \sqrt{\frac{2\pi\hbar^2}{m k_B T}} \tag{4.73}$$

The grand-canonical potential is then

$$J = -k_B T \ln \Xi = -k_B T z e^{\beta\mu} = -k_B T \frac{(2s+1)V}{\Lambda^3} e^{\beta\mu} \tag{4.74}$$

from which we derive the average number of particles

$$\langle N \rangle = -\frac{\partial J}{\partial \mu} = -\beta J = \frac{(2s+1)V}{\Lambda^3} e^{\beta\mu} \tag{4.75}$$

and the pressure

$$P = -\frac{\partial J}{\partial V} = -\frac{J}{V} = k_B T \frac{2s+1}{\Lambda^3} e^{\beta\mu} \tag{4.76}$$

Unsurprisingly, we recover the equation of state

$$PV = \langle N \rangle k_B T \tag{4.77}$$

We may also compute the entropy, recovering the Sackur-Tetrode formula (3.87), and the average energy, recovering the expression (3.84), i.e.,

$$\langle E \rangle = \frac{3}{2} N k_B T \tag{4.78}$$

The reader is invited to do this as an exercise.

## 4.5 The $T - P$ ensemble

This section lies outside of the general pattern of the present chapter, but it presents important results and relations pertaining to systems in which, besides the temperature $T$ and the number of particles $N$, it is not the volume $V$ that is fixed, but the pressure $P$. In that sense, compared to the grand-canonical distribution, this is another generalisation of the canonical distribution, where the system $\mathcal{S}$ is now in contact not with a reservoir imposing $T$ and $\mu$, but with one imposing $T$ and $P$, and with which it can freely exchange energy and volume. The importance of that experimental setup is most apparent in the case of systems in contact with the atmosphere, for instance (e.g., chemical reactions).

### 4.5.1 Thermal and mechanical equilibrium between two systems

Consider two weakly-coupled systems $\mathcal{A}$ and $\mathcal{B}$ separated by a diathermal piston. These systems can therefore exchange volume and energy, but no particles. We assume that $\mathcal{A} \cup \mathcal{B}$ is isolated and rigid, so that

$$E_\mathcal{A} + E_\mathcal{B} = E_0 \qquad V_\mathcal{A} + V_\mathcal{B} = V_0 \tag{4.79}$$

where $E_0$ and $V_0$ are fixed. According to equation (2.39), at equilibrium, the microcanonical pressures and temperatures in both systems should be equal

$$P_\mathcal{A} = P_\mathcal{B} \qquad T_\mathcal{A} = T_\mathcal{B} \tag{4.80}$$

### 4.5.2 System in contact with a reservoir of heat and volume

#### Physical situation

Now, assume $\mathcal{B}$ to be much larger than $\mathcal{A}$, and let us rename the two systems, $\mathcal{A}$ as $\mathcal{S}$ and $\mathcal{B}$ as $\mathcal{R}$ so that it is clear that the latter is now a reservoir of energy and volume, imposing its temperature $T$ and pressure $P$ to the system $\mathcal{S}$. Indeed, with $\mathcal{R}$ much larger than $\mathcal{S}$, the energy and volume variations of the reservoir due to exchanges with the system do not alter its temperature and pressure.

### $T-P$ distribution

We now ask what the probability $w_\ell(V)\mathrm{d}V$ is that the system $\mathcal{S}$ be in microstate $|\ell\rangle$, with energy $E_\ell$ and a volume between $V$ and $V+\mathrm{d}V$. We note that the volume being a continuous variable, this is a bit different to what we did in the grand-canonical approach, where particle numbers are discrete variables, but the principles are the same. This probability must be proportional to the number of microstates of the reservoir whose energy and volume are adequate, i.e.

$$w_\ell(V)\mathrm{d}V \propto \Omega_{\mathcal{R}}\left[E_0-E_\ell, V_0-V\right]\mathrm{d}V \propto \exp\left\{\frac{S_{\mathcal{R}}\left[E_0-E_\ell, V_0-V\right]}{k_B}\right\}\mathrm{d}V \tag{4.81}$$

introducing the microcanonical entropy of the reservoir, which may be expanded to the first order as

$$S_{\mathcal{R}}\left[E_0-E_\ell, V_0-V\right] = S_{\mathcal{R}}\left(E_0, V_0\right) - E_\ell\frac{\partial S_{\mathcal{R}}}{\partial E_{\mathcal{R}}}(E_0, V_0) - V\frac{\partial S_{\mathcal{R}}}{\partial V_{\mathcal{R}}}(E_0, V_0) \tag{4.82}$$

from which we get an expression involving the reservoir's temperature and pressure

$$S_{\mathcal{R}}\left[E_0-E_\ell, V_0-V\right] = S_{\mathcal{R}}\left(E_0, V_0\right) - \frac{E_\ell}{T} - \frac{PV}{T} \tag{4.83}$$

and therefore the expression of the probabilities for the **$T-P$ distribution**

$$w_\ell(V)\mathrm{d}V = \frac{\mathrm{d}V}{\tilde{Z}}\exp\left[-\frac{E_\ell+PV}{k_BT}\right] \tag{4.84}$$

where we have introduced the **$T-P$ partition function**

$$\tilde{Z}(T,P,N) = \int_0^\infty \mathrm{d}V \sum_{|\ell\rangle}\exp\left[-\frac{E_\ell+PV}{k_BT}\right] \tag{4.85}$$

It should be recalled that the energies $E_\ell$ of the microstates of the system may depend on its volume, i.e., $E_\ell(V)$, and we note that this partition function involves not only a discrete sum on these microstates but also an integral on the system's volume $V$.

## 4.5.3 Thermodynamic relations

### Gibbs free-energy

Similarly to the definition of the Helmholtz free-energy $F$ from the canonical partition function $Z$, or to that of the grand-canonical potential $J$ from the grand-canonical partition function $\Xi$, we define the **Gibbs free-energy** $G$ from the $T-P$ partition function

$$G(T,P,N) = -k_BT\ln\tilde{Z} \tag{4.86}$$

### Average volume and energy

There again very similarly to what we have done already in other statistical ensembles, we may compute average values of internal variables, in particular the average volume and average energy as

$$\langle V\rangle = \int_0^\infty V\sum_{|\ell\rangle} w_\ell(V)\mathrm{d}V \qquad\qquad \langle E\rangle = \int_0^\infty \sum_{|\ell\rangle} E_\ell(V)w_\ell(V)\mathrm{d}V \tag{4.87}$$

These can both be computed from the partition function or Gibbs free-energy

$$\langle V\rangle = -k_B T\frac{\partial \ln \tilde{Z}}{\partial P} = \frac{\partial G}{\partial P} \qquad \langle E\rangle = -\frac{\partial \ln \tilde{Z}}{\partial \beta} + \frac{P}{\beta}\frac{\partial \ln \tilde{Z}}{\partial P} = G + \beta\frac{\partial G}{\partial \beta} - P\frac{\partial G}{\partial P} \quad (4.88)$$

We leave it to the reader to demonstrate these relations.

### Entropy

The entropy is computed from the general statistical formula (1.15), adapted to the fact that an integration over the continuous variable $V$ is necessary,

$$S = -k_B \int_0^\infty \mathrm{d}V \sum_{|\ell\rangle} w_\ell(V) \ln\left[w_\ell(V)\right] \quad (4.89)$$

We leave it to the reader to perform this computation to show that

$$S = k_B\left[\ln \tilde{Z} - \beta\frac{\partial \ln \tilde{Z}}{\partial \beta}\right] = -\frac{\partial G}{\partial T} = -\frac{G}{T} + \frac{\langle E\rangle}{T} + \frac{P\langle V\rangle}{T} \quad (4.90)$$

The latter relation is more usually written in a form that is more familiar from classical thermodynamics

$$G = \langle E\rangle + P\langle V\rangle - TS \quad (4.91)$$

### Thermodynamic limit

If the system $\mathcal{S}$ is macroscopic, the fluctuations of the internal variables are negligible, and these may be simply identified with their average or most probable values, so that

$$G = E + PV - TS = F + PV \quad (4.92)$$

where $F$ is the Helmholtz free-energy. In the case of a simple fluid, entropy only depends on three variables, $E$, $V$, and $N$ and since we then have

$$\mathrm{d}S = \frac{1}{T}\mathrm{d}E + \frac{P}{T}\mathrm{d}V - \frac{\mu}{T}\mathrm{d}N \quad (4.93)$$

the differential of the Gibbs free-energy is straightforward to obtain from $G = E + PV - TS$, i.e.

$$\mathrm{d}G = -S\mathrm{d}T + V\mathrm{d}P + \mu\mathrm{d}N \quad (4.94)$$

The chemical potential appears as the partial derivative of $G$ with respect to the number of particles,

$$\mu = \left(\frac{\partial G}{\partial N}\right)_{T,P} \quad (4.95)$$

and since it is an intensive quantity, it can only depend on intensive parameters $T$ and $P$, and not on the sole extensive one $N$, so $\mu = \mu(T, P)$ and the Gibbs free-energy has a simple expression

$$G(T,P,N) = N\mu(T,P) \tag{4.96}$$

This can also be shown from the fact that $G$ is a homogeneous function of order one in $N$ and of order zero in $T$ and $P$, i.e., $G(T,P,\alpha N) = \alpha G(T,P,N)$ and using the same computation we used to show that the grand-canonical potential for a simple fluid is $J = -PV$.

### 4.5.4 Spontaneous evolution towards equilibrium

Imagine that in the system $\mathcal{S}$, whose temperature and pressure are imposed by the reservoir $\mathcal{R}$, some (other) constraint is relaxed. As was the case for the microcanonical case, where such a relaxation led to an increase of the entropy $\mathrm{d}S \geqslant 0$, in the canonical case, where such a relaxation led to a decrease of the Helmholtz free-energy $\mathrm{d}F \leqslant 0$, and in the grand-canonical case, where such a relaxation led to a decrease of the grand-canonical potential $\mathrm{d}J \leqslant 0$, in the $T-P$ case, this relaxation will lead to a decrease of the Gibbs free-energy, towards a minimum that is compatible with the remaining constraints,

$$\mathrm{d}G \leqslant 0 \tag{4.97}$$

This concludes our study of the various statistical ensembles of interest. In each of these, we exhibited a **thermodynamic potential** that reaches an extremum at equilibrium :

— The entropy $S$ is maximum in the microcanonical approach ($E$, $V$, $N$ imposed)
— The Helmholtz free-energy $F$ is minimum in the canonical approach ($T$, $V$, $N$ imposed)
— The grand-canonical potential $J$ is minimum in the grand-canonical approach ($T$, $V$, $\mu$ imposed)
— The Gibbs free-energy $G$ is minimum in the $T$-$P$ approach ($T$, $P$, $N$ imposed)

Of course, in systems where other experimental constraints are imposed, still another statistical ensemble may be relevant. The method, however, will be very similar, and the thermodynamic limit for macroscopic systems ensures that the results will be identical.

# 5

# Ideal fermionic gas

With this chapter, we start discussing the properties of **quantum ideal gases**. A quantum ideal gas is a system of identical, indistinguishable particles whose interactions among themselves are ignored (that is what "ideal" stands for). The indistinguishability requires a treatment taking into account quantum effects depending on whether these particles are fermions or bosons, as explained in the previous chapter. In this chapter we discuss a **fermionic ideal gas**, while the **bosonic** and **photonic** ideal gases will be treated in following chapters. As explained earlier, the computations are made much easier by placing ourselves in the grand-canonical ensemble, so we fix $T$, $V$, and $\mu$. Using the thermodynamic limit, we may then interpret these results in the canonical framework.

## 5.1 General properties

### 5.1.1 The Fermi factor

For an individual state $|\lambda\rangle$, with energy $\epsilon_\lambda$, the **average occupation number** is given by the Fermi-Dirac distribution that we found in the previous chapter

$$\langle N_\lambda \rangle^F = \frac{1}{e^{\beta(\epsilon_\lambda - \mu)} + 1} \tag{5.1}$$

By introducing the density of states, we shall work out the various thermodynamic quantities through integrals over the energy $\epsilon$, rather than through discrete sums over states. Consequently, it makes sense to introduce the **Fermi factor**

$$N^F(\epsilon; T, \mu) = \frac{1}{e^{\beta(\epsilon - \mu)} + 1} \tag{5.2}$$

In the following we may omit the dependence on $T$ and $\mu$, and consider it implicitly. As expected, we have $0 \leqslant N^F \leqslant 1$. This factor, naturally seen as a function of $\epsilon$, is symmetric about the point $(\mu, 1/2)$ which means that

$$N^F(2\mu - \epsilon) = 1 - N^F(\epsilon) \tag{5.3}$$

The shape of the function is shown in Fig. 5.1. There are two limiting regimes

$$\lim_{\epsilon \ll \mu} N^F(\epsilon; T, \mu) = 1 \qquad \lim_{\epsilon \gg \mu} N^F(\epsilon; T, \mu) = 0 \tag{5.4}$$

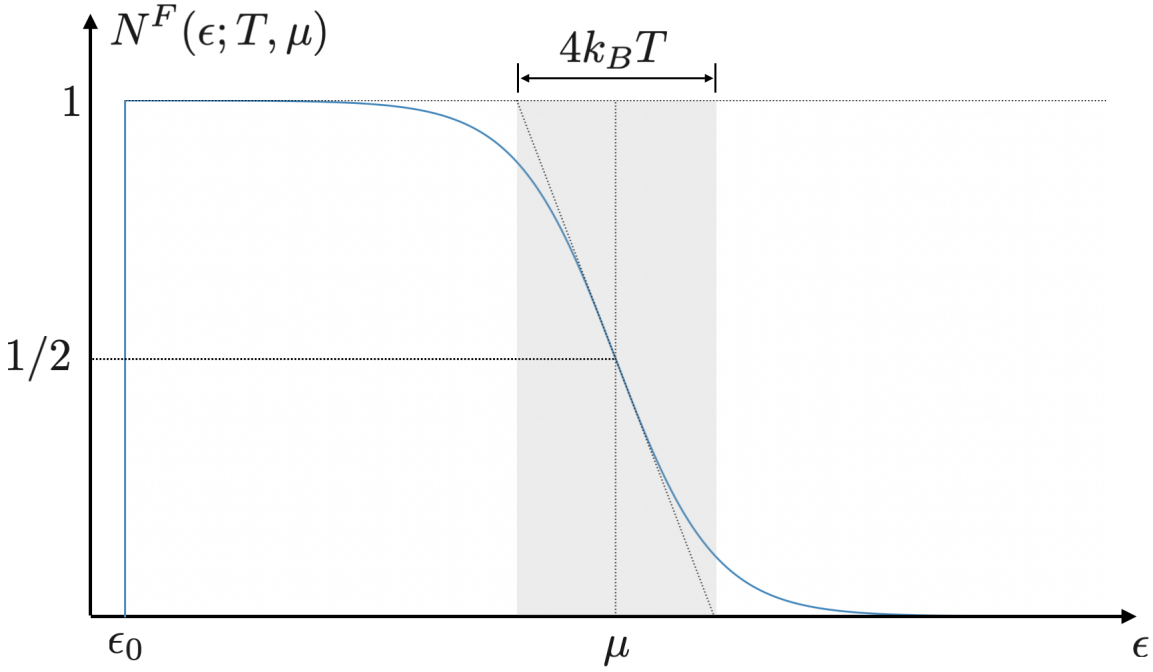



FIGURE 5.1 – The Fermi factor

If $T$ is fixed and $\mu$ varies, the whole shape is simply shifted horizontally, without any deformation. If $\mu$ is fixed and $T$ varies, however, the symmetry point remains fixed, but the shape changes, approaching a step function as $T \to 0$. In that limit, $N^F(\epsilon;0,\mu) = 1$ for $\epsilon < \mu$ and $N^F(\epsilon;0,\mu) = 0$ for $\epsilon > \mu$. The speed of the transition between the two limiting regimes mentioned above is inversely proportional to the temperature, since we have (the demonstration is left to the reader as an exercise)

$$\frac{\partial N^F}{\partial \epsilon}(\mu) = -\frac{1}{4k_BT} \tag{5.5}$$

The transition itself is therefore occurring over a range of energies $\Delta\epsilon \sim 4k_BT$. Various Fermi factors corresponding to different temperatures are shown in Fig. 5.2.

### 5.1.2 Physical quantities

The expressions for the average number of particles and average energy in the grand-canonical approach are given in equations (4.56), and may be converted into integrals over the energy using the density of states. For the average number of particles we then have

$$\langle N\rangle = \sum_{|\lambda\rangle}\langle N_\lambda\rangle^F = \int_{\epsilon_0}^{\infty}\rho(\epsilon)N^F(\epsilon)\mathrm{d}\epsilon = \int_{\epsilon_0}^{\infty}\rho(\epsilon)\frac{1}{e^{\beta(\epsilon-\mu)}+1}\mathrm{d}\epsilon \tag{5.6}$$

and for the average energy we have, similarly,

$$\langle E\rangle = \sum_{|\lambda\rangle}\langle N_\lambda\rangle^F\epsilon_\lambda = \int_{\epsilon_0}^{\infty}\rho(\epsilon)\epsilon N^F(\epsilon)\mathrm{d}\epsilon = \int_{\epsilon_0}^{\infty}\rho(\epsilon)\frac{\epsilon}{e^{\beta(\epsilon-\mu)}+1}\mathrm{d}\epsilon \tag{5.7}$$

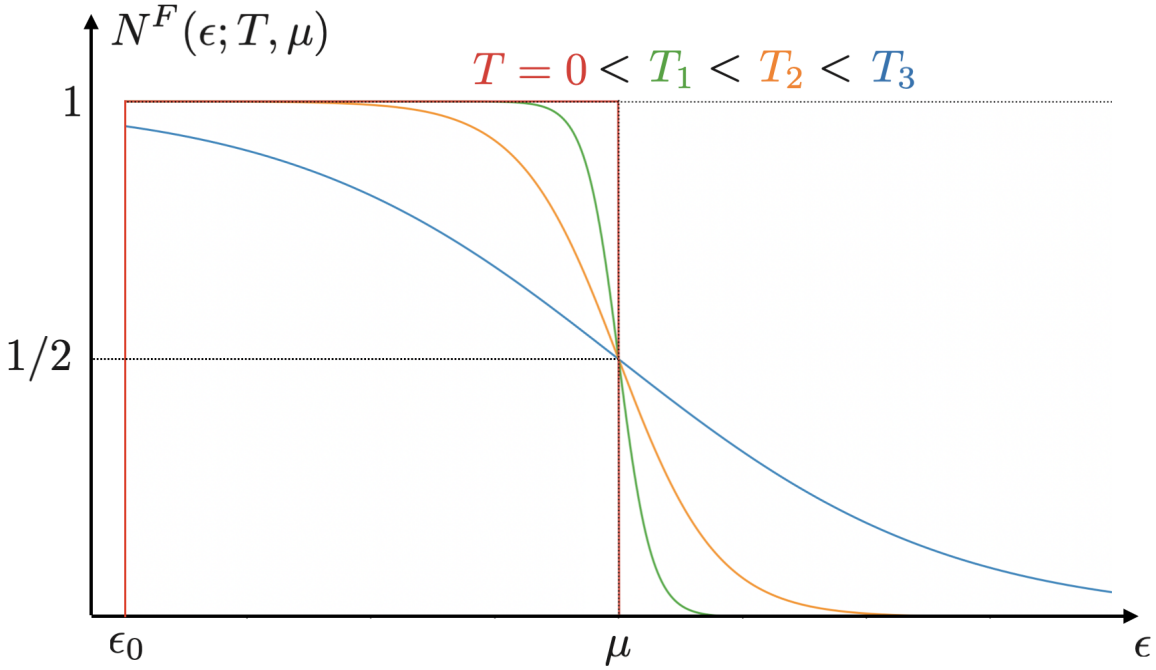



FIGURE 5.2 – Fermi factors for various temperatures

The grand-canonical potential may be computed using a similar expression [1]

$$J = -k_B T \sum_{|\lambda\rangle} \ln \xi_\lambda^F = k_B T \sum_{|\lambda\rangle} \ln\left[1 - \langle N_\lambda\rangle^F\right] = k_B T \int_{\epsilon_0}^{\infty} \rho(\epsilon) \ln\left[1 - N^F(\epsilon)\right] \mathrm{d}\epsilon \tag{5.8}$$

Other thermodynamic quantities may be computed in the same way. The reader is encouraged to try and compute the entropy $S$ as an exercise.

### 5.1.3 Canonical interpretation

Now, all of this was computed in the grand-canonical ensemble, but oftentimes we are rather dealing with systems with a fixed number of particles, so we should work in the canonical ensemble. The thermodynamic limit ensures that for large enough systems, the two descriptions are equivalent, due to the very small range of fluctuations possible for internal variables, including $N$. Consequently, if we work in the grand-canonical ensemble, fixing $T$, $V$, and $\mu$, we may compute the average number $\langle N\rangle(T,V,\mu)$ and the relationship between these quantities may be (at least theoretically) inverted to yield $\mu(T,V,N)$ which would be the (canonical) chemical potential computed for this system in the canonical ensemble, for the given $N$. To summarize, seen in the canonical framework, the equation

$$N = \int_{\epsilon_0}^{\infty} \rho(\epsilon) \frac{1}{e^{\beta(\epsilon-\mu)}+1} \mathrm{d}\epsilon \tag{5.9}$$

1. We recall that for a Fermi-Dirac distribution, $1 - \langle N_\lambda\rangle^F = 1/\xi_\lambda^F$.

is an implicit equation in $\mu$. Once it is solved, the (average) energy may be computed from

$$\langle E\rangle = \int_{\epsilon_0}^{\infty} \rho(\epsilon)\frac{\epsilon}{e^{\beta(\epsilon-\mu)}+1}\mathrm{d}\epsilon \tag{5.10}$$

by inserting this value $\mu(T,V,N)$ of the chemical potential.

### 5.1.4 Example of the three-dimensional ideal gas

The above expressions have not yet made use of the hypotheses regarding the system, keeping the density of states $\rho(\epsilon)$ and the ground energy $\epsilon_0$ unspecified. Now, we focus on the example of an ideal gas of fermions in a three-dimensional box of volume $V$. For simplicity, we assume that the particles have no internal structure (no electronic levels, no rotation, and no vibration), except for a spin $s$, that they are free, i.e., do not interact with the outside world, which means that they are not subject to any external field, and that they are non-relativistic. In that case, the energy $\epsilon$ and the density of states $\rho(\epsilon)$ are given by

$$\epsilon = \frac{p^2}{2m} = \frac{\hbar^2k^2}{2m} \qquad\qquad \rho(\epsilon) = \frac{(2s+1)Vm^{3/2}}{\sqrt{2}\pi^2\hbar^3}\sqrt{\epsilon} = AV\sqrt{\epsilon} \tag{5.11}$$

where $m$ is the mass of a particle, and $s$ its spin. In the non-relativistic regime, the ground-level energy is $\epsilon_0 = 0$. We have introduced the following constant $A$ for notational simplicity,

$$A = \frac{(2s+1)m^{3/2}}{\sqrt{2}\pi^2\hbar^3} \tag{5.12}$$

#### Determination of the canonical chemical potential

The chemical potential is determined by writing the implicit equation giving the number of particles

$$\langle N\rangle = \int_0^{\infty} \frac{AV\sqrt{\epsilon}}{e^{\beta(\epsilon-\mu)}+1}\mathrm{d}\epsilon \tag{5.13}$$

We introduce the fugacity $\varphi = e^{\beta\mu} > 0$ and change to the variable $x = \beta\epsilon$, so that

$$\frac{\langle N\rangle}{AV\left(k_BT\right)^{3/2}} = \int_0^{\infty} \frac{\sqrt{x}}{\frac{e^x}{\varphi}+1}\mathrm{d}x = I_F(\varphi) \tag{5.14}$$

The function $I_F(\varphi)$ is represented in Fig. 5.3. It is monotonously increasing from $I_F(0) = 0$ to infinity when $\varphi \to \infty$, which ensures that for a given number of particles $N$ in the canonical approach, there is a single solution $\varphi_0$ such that $N/AV\left(k_BT\right)^{3/2} = I_F(\varphi_0)$, ensuring therefore the existence and uniqueness of the chemical potential $\mu(T,V,N) = k_BT\ln\varphi_0$, as explained graphically in Fig. 5.3.

#### Average energy and grand-canonical potential

With this solution $\mu(T,V,N)$, the energy[2] reads

$$E = \int_0^{\infty} \frac{AV\epsilon^{3/2}}{e^{\beta(\epsilon-\mu)}+1}\mathrm{d}\epsilon \tag{5.15}$$

and the grand-canonical potential is

$$J = k_BT\int_0^{\infty} AV\epsilon^{1/2}\ln\left[1-N^F(\epsilon)\right]\mathrm{d}\epsilon = -k_BT\int_0^{\infty} AV\epsilon^{1/2}\ln\left[1+e^{-\beta(\epsilon-\mu)}\right]\mathrm{d}\epsilon \tag{5.16}$$

2. In the thermodynamic limit, we may write $E$ without specifying that it is an average.

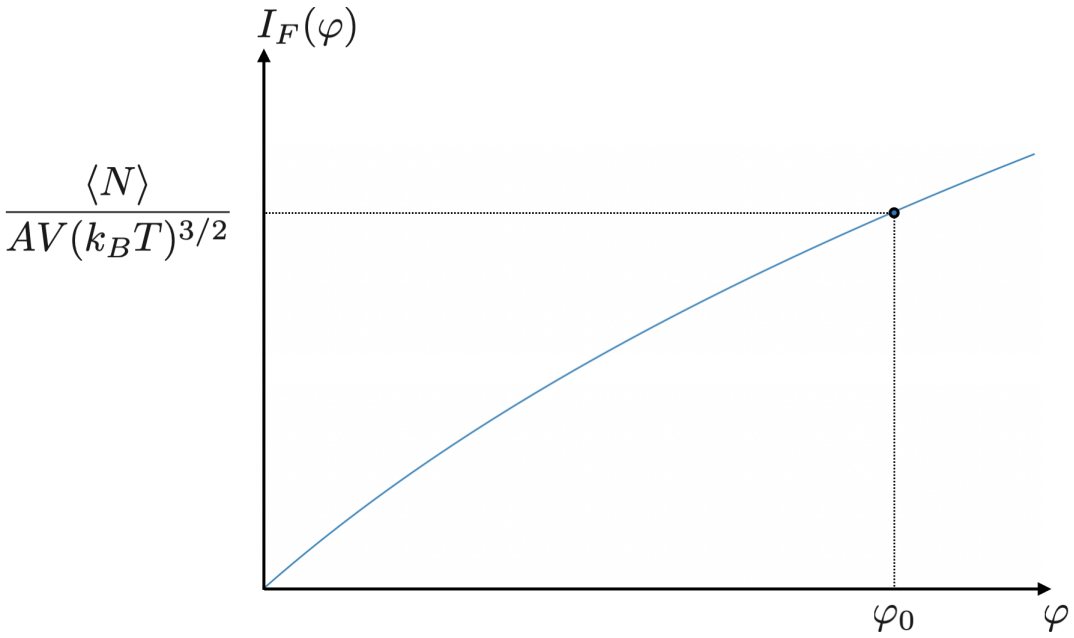



FIGURE 5.3 – The function $I_F(\varphi)$ and graphical determination of the single solution $\varphi_0$.

Through an integration by parts, we have

$$J = -k_BT\left\{\left[\frac{2}{3}AV\epsilon^{3/2}\ln\left[1+e^{-\beta(\epsilon-\mu)}\right]\right]_0^\infty + \frac{2\beta}{3}\int_0^\infty AV\epsilon^{3/2}\frac{e^{-\beta(\epsilon-\mu)}}{1+e^{-\beta(\epsilon-\mu)}}\mathrm{d}\epsilon\right\} \tag{5.17}$$

The fully integrated term is null [3], and the integral term simplifies to

$$J = -\frac{2}{3}AV\int_0^\infty \frac{\epsilon^{3/2}}{e^{\beta(\epsilon-\mu)}+1}\mathrm{d}\epsilon \tag{5.18}$$

so that for any temperature, we have the simple relation

$$J = -\frac{2}{3}E \tag{5.19}$$

**Entropy and pressure**

From the grand-canonical potential, we may derive expressions for the entropy and the pressure

$$S = -\left(\frac{\partial J}{\partial T}\right)_{V,\mu} \qquad P = -\left(\frac{\partial J}{\partial V}\right)_{T,\mu} \tag{5.20}$$

For the entropy, we shall see later that it tends to zero in the limit of zero temperature. For now, we focus on the pressure, which is especially simple to compute, considering that $J$ is linear in $V$, so that

3. The term for $\epsilon \to 0$ is obviously null, and that for $\epsilon \to +\infty$ is also null by considering that $\ln(1+x) \approx x$ when $x = e^{-\beta(\epsilon-\mu)} \to 0$ in this limit.

$$P = -\frac{J}{V} = \frac{2}{3}\frac{E}{V} \tag{5.21}$$

The relation $PV = (2/3)E$ is therefore valid for an ideal, non-relativistic fermionic gas at any temperature [4]. We note that the numerical factor $2/3$ is directly related to the scaling of the density of states $\rho(\epsilon) \propto \epsilon^{1/2}$.

### Classical limit

The classical limit is obtained when the average occupation number is very low, so when $e^{\beta(\epsilon-\mu)} \gg 1$. In that case the implicit equation giving the chemical potential becomes

$$N = \int_0^\infty AV\epsilon^{1/2}e^{-\beta(\epsilon-\mu)}\mathrm{d}\epsilon = AVe^{\beta\mu}(k_BT)^{3/2}\int_0^\infty x^{1/2}e^{-x}\mathrm{d}x \tag{5.22}$$

The integral is found using Euler's Gamma function

$$\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}\mathrm{d}x \qquad \Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2} \qquad \Gamma\left(\frac{5}{2}\right) = \frac{3}{4}\sqrt{\pi} \tag{5.23}$$

We therefore have the implicit equation

$$N = AV\frac{\sqrt{\pi}}{2}e^{\beta\mu}(k_BT)^{3/2} \tag{5.24}$$

and for the record, the canonical chemical potential that is the solution of the implicit equation is, in this classical limit,

$$\mu = k_BT\ln\left[\frac{2N}{AV\sqrt{\pi}(k_BT)^{3/2}}\right] \tag{5.25}$$

The energy of the system is then given, in this limit, by

$$E = \int_0^\infty AV\epsilon^{3/2}e^{-\beta(\epsilon-\mu)}\mathrm{d}\epsilon = AVe^{\beta\mu}(k_BT)^{5/2}\int_0^\infty x^{3/2}e^{-x}\mathrm{d}x = \frac{3}{4}\sqrt{\pi}AVe^{\beta\mu}(k_BT)^{5/2} \tag{5.26}$$

and may be directly and simply related to the number of particles as

$$\frac{E}{N} = \frac{3}{2}k_BT \tag{5.27}$$

Since $E = (3/2)PV$, we recover the usual equation of state of the classical ideal gas

$$PV = Nk_BT \tag{5.28}$$

It is important to stress that the two equations (5.27) and (5.28) are only valid in the classical limit, but that the relation $2E = 3PV$ is valid in any regime for this system. An interesting point to make in this respect is precisely to discuss the range of temperatures for which the classical approximation is valid. As mentioned, it is correct if $e^{\beta(\epsilon-\mu)} \gg 1$, so it should be true for the ground level, i.e., $e^{\beta\mu} \ll 1$, which, given the above expression for the chemical potential, amounts to [5]

$$T \gg \frac{1}{k_B}\left(\frac{N}{AV}\right)^{2/3} \tag{5.29}$$

4. For an ultra-relativistic fermionic gas, we find a different relation, $P = E/(3V)$ as for the photon gas (Chapter 7).
5. We ignore the unimportant numerical factor of order one.

This is equivalent, up to unimportant numerical factors, to the condition $d = (V/N)^{1/3} \gg \Lambda$ already determined [6]. In this limit, a classical treatment is appropriate. In the opposite limit, at low temperatures, a treatment taking into account **quantum degeneracies** is warranted, but results may be obtained analytically as we shall see shortly. In the intermediate regime, a fully numerical treatment is necessary.

## 5.2 Low-temperature limit

### 5.2.1 The case $T = 0$ and Fermi quantities

We remain in the non-relativistic regime, but now assume that the temperature is very low, i.e.,

$$T \ll \frac{1}{k_B}\left(\frac{N}{AV}\right)^{2/3} \tag{5.32}$$

and we shall take in this section that $T \to 0$. The solution of the implicit equation for $\mu$ is now

$$\mu_0(V,N) = \lim_{T\to 0}\mu(T,V,N) \tag{5.33}$$

and is quite easily determined, since the Fermi factor at $T = 0$ is a simple step function

$$N = \int_0^{\mu_0} AV\epsilon^{1/2}\mathrm{d}\epsilon = \frac{2}{3}AV\mu_0^{3/2} \tag{5.34}$$

and so we have the chemical potential at $T = 0$ with the following formula, which gives the so-called **Fermi level** or **Fermi energy**, that is the highest energy level being populated at zero-temperature [7]

$$\epsilon_F = \mu_0 = \frac{\hbar^2}{2m}\left(\frac{6\pi^2}{2s+1}\frac{N}{V}\right)^{2/3} \tag{5.35}$$

For an ideal gas, the energy is directly related to the momentum and the wavevector, so we may define the **Fermi momentum** and the **Fermi wavenumber** via

$$p_F = \sqrt{2m\mu_0} = \hbar\left(\frac{6\pi^2}{2s+1}\frac{N}{V}\right)^{1/3} \qquad k_F = \left(\frac{6\pi^2}{2s+1}\frac{N}{V}\right)^{1/3} \tag{5.36}$$

The tip of the wavevectors in $\vec{k}$-space that correspond to the highest-energy states describes a surface, called the **Fermi surface**. At $T = 0$, all states interior to that surface are populated, and all states outside of it are empty. For an ideal gas (free particles), that surface is a sphere, called the **Fermi sphere**. For systems including interactions, such as electrons in a metal, the Fermi surfaces will have more complex shapes. We may also define a **Fermi temperature** [8] through

6. Indeed, the condition above is

$$k_B T \gg \left(\frac{N}{AV}\right)^{2/3} \approx \left(\frac{N\hbar^3}{m^{3/2}V}\right)^{2/3} = \frac{\hbar^2}{m}\left(\frac{N}{V}\right)^{2/3} \tag{5.30}$$

and so is equivalent to

$$\left(\frac{V}{N}\right)^{1/3} \gg \frac{\hbar}{\sqrt{mk_BT}} \approx \Lambda \tag{5.31}$$

7. We leave it to the reader to check that the numerical factors are correct.

8. Recalling the expression of the constant $A$, the condition for a low-temperature limit reads

$$T \ll \frac{1}{k_B}\left(\frac{N}{V}\right)^{2/3}\left[\frac{\sqrt{2}\pi^2\hbar^3}{(2s+1)m^{3/2}}\right]^{2/3} = \left(\frac{2}{3}\right)^{2/3} T_F \approx T_F \tag{5.37}$$

$$T_F = \frac{\mu_0}{k_B} = \frac{\hbar^2}{2mk_B}\left(\frac{6\pi^2}{2s+1}\frac{N}{V}\right)^{2/3} \tag{5.38}$$

In the limit $T \to 0$, the energy reads

$$E(T=0,V,N) = E_0 = \int_0^{\mu_0} AV\epsilon^{3/2}\mathrm{d}\epsilon = \frac{2}{5}AV\mu_0^{5/2} = \frac{3}{5}N\mu_0 = \frac{3}{5}Nk_BT_F \tag{5.39}$$

As for the pressure, it reads

$$P(T=0,V,N) = P_0 = -\frac{J_0}{V} = \frac{2}{3}\frac{E_0}{V} = \frac{4}{15}A\mu_0^{5/2} = \frac{2}{5}\frac{Nk_BT_F}{V} \tag{5.40}$$

Summarizing, we have that **the energy and the pressure are strictly positive**

$$E(T=0,V,N) = \frac{3}{5}Nk_BT_F \qquad P(T=0,V,N) = \frac{2}{5}\frac{Nk_BT_F}{V} \tag{5.41}$$

They do not vanish at zero temperature, unlike what a classical treatment would have (since $E \propto T$ and $P \propto T$ in that case). This is a purely quantum effect, due to the Pauli exclusion principle. It is this very pressure that supports white dwarf stars against gravity.

### 5.2.2 Low-temperature behaviour ($0 < T \ll T_F$)

#### Sommerfeld expansion

At low, but non-zero temperatures, we may compute the previous quantities (chemical potential and energy in particular) to the next leading order in $T/T_F$. This is called the **Sommerfeld expansion**. It involves computing integrals of the form

$$g(T,\mu) = \int_0^\infty \mathrm{d}\epsilon f(\epsilon)N^F(\epsilon;T,\mu) \tag{5.42}$$

for sufficiently well-behaved functions $f$. The expressions above for the number of particles, energy, grand-canonical potential and the like, are all of this form. It may be shown, and should be done as an exercise, that in the limit $T \to 0$, we have

$$g(T,\mu) = g(0,\mu) + \frac{\pi^2}{6}\left(k_BT\right)^2 f'(\mu) + \mathcal{O}\left(T^4\right) \tag{5.43}$$

#### Chemical potential

Applying this to the number of particles, with therefore $f(\epsilon) = AV\epsilon^{1/2}$, we may derive the expression of the chemical potential at low temperature, starting from

$$N(T,\mu) = N(0,\mu) + \frac{\pi^2}{12}\left(k_BT\right)^2\frac{AV}{\mu^{1/2}} + \mathcal{O}\left(T^4\right) \tag{5.44}$$

so that when $T \ll T_F$ we should treat the gas in this fully degenerate approach, while a classical treatment is sufficient in the opposite limit $T \gg T_F$.

The second term on the right-hand side being a correction, we may replace $\mu$ by $\mu_0$ in it [9], and the first term is given by the expression at $T = 0$ already found (5.34), but for $\mu$, not $\mu_0$, so

$$N(T,\mu) = \frac{2}{3}AV\mu^{3/2} + \frac{\pi^2}{12}(k_BT)^2\frac{AV}{\mu_0^{1/2}} + \mathcal{O}\left(T^4\right) \tag{5.45}$$

Now the left-hand side is precisely the imposed number of particles. It is equal to what it would be at zero temperature, so that it can be expressed with the proper relation involving $\mu_0$,

$$\frac{2}{3}AV\mu_0^{3/2} = \frac{2}{3}AV\mu^{3/2} + \frac{\pi^2}{12}(k_BT)^2\frac{AV}{\mu_0^{1/2}} + \mathcal{O}\left(T^4\right) \tag{5.46}$$

This may be inverted to isolate $\mu$, yielding

$$\mu^{3/2} = \mu_0^{3/2} - \frac{\pi^2}{8}\frac{(k_BT)^2}{\mu_0^{1/2}} + \mathcal{O}\left(T^4\right) = \mu_0^{3/2}\left[1 - \frac{\pi^2}{8}\left(\frac{k_BT}{\mu_0}\right)^2 + \mathcal{O}\left(T^4\right)\right] \tag{5.47}$$

where we recognize the Fermi temperature $T_F = \mu_0/k_B$. Finally, we take the power $2/3$ and apply the linear development $(1-x)^\alpha \approx 1 - \alpha x$ to obtain the chemical potential at low temperature

$$\mu(T,V,N) = \mu_0\left[1 - \frac{\pi^2}{12}\left(\frac{T}{T_F}\right)^2 + \mathcal{O}\left(\frac{T^4}{T_F^4}\right)\right] \tag{5.48}$$

We note that $\mu$ decreases as $T$ increases, which is expected considering Fig. 5.3. This expansion is shown as the solid blue line in Fig. 5.4, alongside data points that correspond to a complete numerical determination based on Eq. (5.14).

### Energy

Using now $f(\epsilon) = AV\epsilon^{3/2}$, we may compute the expression of the energy at low temperature, starting from the general expression [10]

$$E(T,\mu) = E(0,\mu) + \frac{\pi^2}{4}AV\mu_0^{1/2}(k_BT)^2 + \mathcal{O}\left(T^4\right) \tag{5.49}$$

in which the left-hand-side is just the energy we seek $E$, and we replace the first term on the right-hand side by the expression (5.39) for $\mu$

$$E = \frac{2}{5}AV\mu^{5/2} + \frac{\pi^2}{4}AV\mu_0^{1/2}(k_BT)^2 + \mathcal{O}\left(T^4\right) \tag{5.50}$$

We then use the above expression for $\mu$ at low temperature

$$E = \frac{2}{5}AV\mu_0^{5/2}\left[1 - \frac{\pi^2}{12}\left(\frac{T}{T_F}\right)^2 + \mathcal{O}\left(\frac{T^4}{T_F^4}\right)\right]^{5/2} + \frac{\pi^2}{4}AV\mu_0^{1/2}(k_BT)^2 + \mathcal{O}\left(T^4\right) \tag{5.51}$$

9. The reader may want to check that inserting $\mu = \mu_0(1+x)$ with $|x| \ll 1$ in **both** terms on the right-hand side and developing, one gets, as expected from Eq. (5.48),

$$x = -\frac{\pi^2}{12}\left(\frac{T}{T_F}\right)^2$$

10. In the second term on the right-hand side we use $\mu_0$ instead of $\mu$ because the correction would be of order four in $T/T_F$ and absorbed in the $\mathcal{O}\left(T^4\right)$.

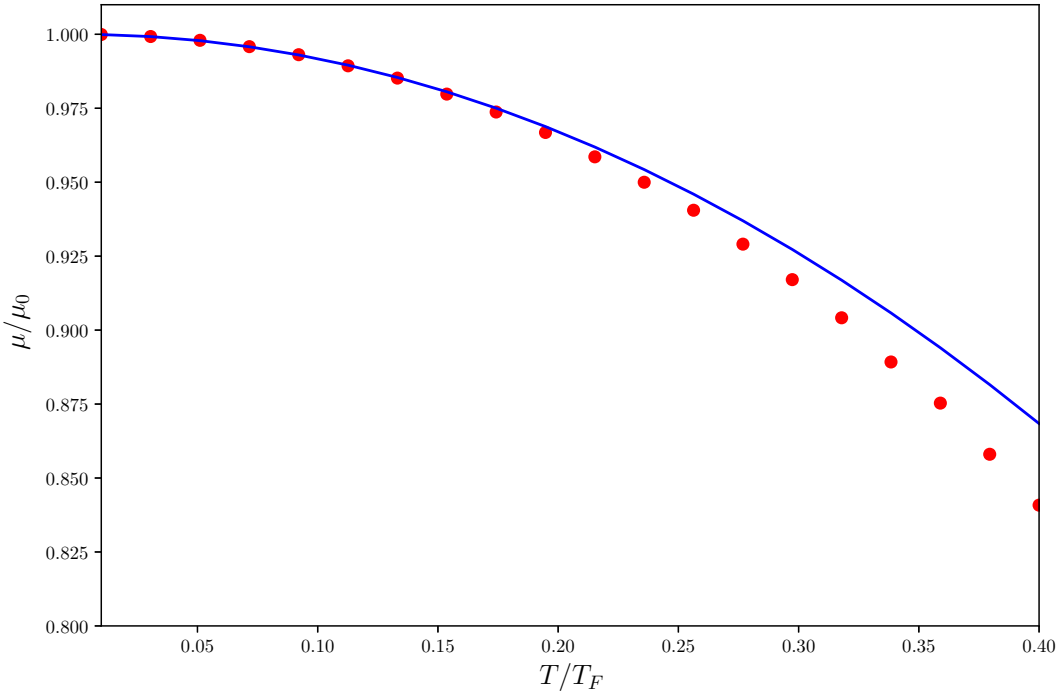



FIGURE 5.4 – Evolution of the chemical potential $\mu$, normalized to the zero-temperature value $\mu_0$ (i.e. the Fermi energy $\epsilon_F$), as a function of the normalized temperature $T/T_F$. The red points correspond to an "exact" numerical determination based on Eq. (5.14), while the solid blue line corresponds to the Sommerfeld expansion (5.48).

in which we can use a development of the expression in square brackets, to the same order, and factor out the zero-temperature energy $E_0$, recalling that $\mu_0 = k_B T_F$,

$$E = \frac{2}{5} A V \mu_0^{5/2} \left\{ \left[ 1 - \frac{5\pi^2}{24} \left( \frac{T}{T_F} \right)^2 \right] + \frac{5\pi^2}{8} \left( \frac{T}{T_F} \right)^2 + \mathcal{O}\left(T^4\right) \right\} \tag{5.52}$$

leading *in fine* to the expression of the energy in the low-temperature regime

$$E(T,V,N) = E_0 \left[ 1 + \frac{5\pi^2}{12} \left( \frac{T}{T_F} \right)^2 + \mathcal{O}\left( \frac{T^4}{T_F^4} \right) \right] \tag{5.53}$$

Unsurprisingly, the energy increases from $E_0$ as the temperature rises.

### Heat capacity

From the expression of the energy, we see that the heat capacity at low temperatures is linear in $T$, so it goes down to zero when $T \to 0$, with [11]

11. We have

$$C_v = \left( \frac{\partial E}{\partial T} \right)_{V,N} = E_0 \frac{5\pi^2}{6} \frac{T}{T_F^2} = \frac{3}{5} N k_B T_F \times \frac{5\pi^2}{6} \frac{T}{T_F^2} = \frac{\pi^2}{2} N k_B \frac{T}{T_F} \tag{5.54}$$

$$C_v = \left(\frac{\partial E}{\partial T}\right)_{V,N} = \eta T \qquad \eta = \frac{\pi^2}{2}\frac{Nk_B}{T_F} \tag{5.55}$$

This is in contrast to the classical regime, where the energy is linear in $T$, $E \sim Nk_BT$, so that the heat capacity is a constant, $C_v \sim Nk_B$. An interesting interpretation of this is the following. At $T = 0$, all states are filled up to $\mu_0$, and at low but non-zero temperatures, the thermostat allows fermions to gain an energy $\sim k_BT$. However, only those fermions lying less than $\sim k_BT$ below the Fermi level may benefit from this, since the lower-lying ones cannot be excited to a level that is already occupied. In effect, the effective number of fermions that can be excited is about

$$N_{\text{eff}} \sim \rho(\mu_0) \times k_BT \sim \frac{Nk_BT}{\mu_0} \tag{5.56}$$

since $\rho(\mu_0) = AV\mu_0^{1/2}$ and $N \sim AV\mu_0^{3/2}$ so that $\rho(\mu_0) \sim N/\mu_0$. The energy of the system is then that at $T = 0$ augmented by $k_BT$ for these potentially excited fermions

$$E \sim E_0 + N_{\text{eff}}k_BT = E_0 + \frac{N}{\mu_0}\left(k_BT\right)^2 = E_0 + \frac{Nk_BT^2}{T_F} \tag{5.57}$$

leading to the correct order of magnitude for the heat capacity

$$C_v \sim Nk_B\frac{T}{T_F} \tag{5.58}$$

**Entropy**

To determine the entropy, we start from its expression derived from the grand-canonical potential

$$S = -\left(\frac{\partial J}{\partial T}\right)_{V,\mu} \tag{5.59}$$

where this potential is directly proportional to the energy, as already mentioned

$$J = -\frac{2}{3}E \tag{5.60}$$

A reckless computation would have us write that $S = (2/3)C_v$, but the derivation in Eq. (5.59) is performed at $V$ and $\mu$ constant, while the heat capacity is computed as a derivative at $V$ and $N$ constant. Therefore, to perform the adequate derivation, we need to express $J$ is terms of the proper grand-canonical variables $(T, V, \mu)$. It is most easily done by using the expression (5.50), in which we take $\mu_0 \approx \mu$ in the second term, since it is a correction,

$$J = -\frac{4}{15}AV\mu^{5/2} - \frac{\pi^2}{6}AV\mu^{1/2}\left(k_BT\right)^2 + \mathcal{O}\left(T^4\right) \tag{5.61}$$

The derivation with respect to $T$ at fixed volume $V$ and chemical potential $\mu$ then yields

$$S = \frac{\pi^2}{3}k_B^2AV\mu^{1/2}T + \mathcal{O}\left(T^3\right) \tag{5.62}$$

To express this is terms of the canonical variables, we simply replace $\mu$ by its expression (5.48), which to that order is just $\mu \approx \mu_0$, and replacing $\mu_0$ by its expression (5.35), leading to

$$S = \frac{\pi^2}{2} N k_B \frac{T}{T_F} + \mathcal{O}\left(\frac{T^3}{T_F^3}\right) \tag{5.63}$$

We note that, to this order, $S = C_v$, and that, as $T \to 0$, the entropy goes down to zero as well. This latter property is the **Nernst principle**, also called **third law of thermodynamics**, which states that the entropy of any system must vanish when the temperature goes to zero [12].

### Equation of state

As an exercise, we leave it to the reader to derive the expressions for the Helmholtz free-energy

$$F(T,V,N) = \frac{3}{5} N k_B T \left[1 - \frac{5\pi^2}{12}\left(\frac{T}{T_F}\right)^2 + \mathcal{O}\left(\frac{T^4}{T_F^4}\right)\right] \tag{5.64}$$

and for the pressure

$$P(T,V,N) = P_0 \left[1 + \frac{5\pi^2}{12}\left(\frac{T}{T_F}\right)^2 + \mathcal{O}\left(\frac{T^4}{T_F^4}\right)\right] \tag{5.65}$$

where the zero-temperature pressure is given by Eq. (5.41)

$$P_0 = \frac{2}{5}\frac{N k_B T_F}{V} \tag{5.66}$$

From the pressure, we may derive the equation of state of the ideal fermionic gas at low temperatures. Explicitly writing out the expression of the zero-temperature pressure $P_0$ and the Fermi temperature as a function of the canonical variables $N$ and $V$, we get

$$P^3 \left(\frac{V}{N}\right)^5 = \frac{36\pi^4}{125(2s+1)^2}\frac{\hbar^6}{m^3}\left[1 + \frac{5\pi^2 m^2 k_B^2 T^2}{\hbar^4}\left(\frac{2s+1}{6\pi^2}\frac{V}{N}\right)^{4/3} + \mathcal{O}\left(T^4\right)\right]^3 \tag{5.67}$$

This equation of state, of the form

$$P^3 \left(\frac{V}{N}\right)^5 = \alpha\left[1 + \beta T^2 \left(\frac{V}{N}\right)^{4/3}\right]^3 \tag{5.68}$$

is indeed quite different from the high-temperature, classical equation of state $PV = N k_B T$ !

### Validity domain

As already stated, the validity domain of the above expressions is that the temperature should be small compared to the Fermi temperature,

$$T \ll T_F = \frac{\hbar^2}{2m k_B}\left(\frac{6\pi^2}{2s+1}\frac{N}{V}\right)^{2/3} \tag{5.69}$$

and in that case these expressions provide first-order corrections to the $T = 0$ case. In the opposite limit $T \gg T_F$, a classical treatment is possible, using the Maxwell-Boltzmann approximation as we did earlier. In between, only a numerical treatment is possible, but we insist that **there is no special behaviour at $T = T_F$, all physical quantities are continuous at this point**.

12. We note that the Sackur-Tetrode formula (3.87) does not obey this principle, but we recall that it applies to a classical ideal gas, i.e., at high temperatures only, and so cannot be expected to hold as $T \to 0$.

## 5.3 Relativistic case

The relativistic treatment is warranted in some cases, when the fermions at the Fermi level reach relativistic speeds, i.e., $p_F \sim mc$. In that case, a number of computations must be undertaken anew.

### 5.3.1 Density of states

We recall that in the classical (non-relativistic) regime, the density of states for a free particle with mass $m$ and spin $s$ in a box of volume $V$ is

$$\rho(\epsilon) = \frac{(2s+1)Vm^{3/2}}{\sqrt{2}\pi^2\hbar^3}\epsilon^{1/2} \tag{5.70}$$

However this is incorrect in the relativistic case. The density of states was computed as the derivative of the cumulative density of states

$$\Phi(\epsilon) = \frac{(2s+1)\sqrt{2}Vm^{3/2}}{3\pi^2\hbar^3}\epsilon^{3/2} \tag{5.71}$$

which itself was computed by counting the number of wavevectors compatible with the constraint that the associated energy was at most $\epsilon$. More generally, the cumulative density of states is then defined by

$$\Phi(\epsilon) = \frac{(2s+1)Vk(\epsilon)^3}{6\pi^2} \tag{5.72}$$

where $k(\epsilon)$ is the relationship between the wavenumber and the energy. In the classical case, we have

$$k(\epsilon) = \frac{\sqrt{2m\epsilon}}{\hbar} \tag{5.73}$$

which yields the above density of states $\rho(\epsilon)$. In the relativistic case, however, the relation is different

$$k(\epsilon) = \frac{1}{\hbar c}\sqrt{\epsilon^2 - m^2c^4} \tag{5.74}$$

and we therefore have the following expressions for the cumulative density of states $\Phi$ and the density of states $\rho$ in the relativistic case

$$\Phi(\epsilon) = \frac{(2s+1)V}{6\pi^2\hbar^3c^3}(\epsilon^2 - m^2c^4)^{3/2} \qquad \rho(\epsilon) = \frac{(2s+1)V}{2\pi^2\hbar^3c^3}(\epsilon^2 - m^2c^4)^{1/2}\epsilon \tag{5.75}$$

### 5.3.2 Fermi quantities at zero temperature

Placing ourselves in the limit $T \to 0$, we may compute the chemical potential $\mu_0(V, N)$ in this limit implicitly, using the equation that gives the total number of particles, and replacing the Fermi factor by the appropriate step function. The one difference with the classical case is that the ground state energy is no longer $\epsilon_0 = 0$, but $\epsilon_0 = mc^2$. We then have

$$N = \int_{\epsilon_0}^{\mu_0} \rho(\epsilon)\mathrm{d}\epsilon = \Phi(\mu_0) - \Phi(mc^2) = \frac{(2s+1)V}{6\pi^2\hbar^3c^3}(\mu_0^2 - m^2c^4)^{3/2} \tag{5.76}$$

The chemical potential, or Fermi energy $\epsilon_F = \mu_0$ is then, inverting this relation,

$$\epsilon_\mathrm{F} = \mu_0 = \sqrt{m^2c^4 + \left(\frac{6\pi^2\hbar^3c^3}{2s+1}\frac{N}{V}\right)^{2/3}} = mc^2\sqrt{1 + \frac{\hbar^2}{m^2c^2}\left(\frac{6\pi^2}{2s+1}\frac{N}{V}\right)^{2/3}} \tag{5.77}$$

From there, we get the Fermi momentum, that has the same expression as in the non-relativistic case,

$$p_{\mathrm{F}} = \frac{1}{c}\sqrt{\epsilon_{\mathrm{F}}^2 - m^2c^4} = \frac{1}{c}\left(\frac{6\pi^2\hbar^3c^3}{2s+1}\frac{N}{V}\right)^{1/3} = \hbar\left(\frac{6\pi^2}{2s+1}\frac{N}{V}\right)^{1/3} \tag{5.78}$$

The definition of the Fermi temperature is a bit different from the non-relativistic case. It is defined such that $k_B T_{\mathrm{F}} = \mu_0(V,N) - \epsilon_0$, i.e., $k_B T_F$ is the energy range between the ground-level and the Fermi level [13]. Instead of the non-relativistic expression (5.38), we now have

$$T_{\mathrm{F}} = \frac{\epsilon_{\mathrm{F}} - mc^2}{k_B} = \frac{mc^2}{k_B}\left[\sqrt{1 + \frac{\hbar^2}{m^2c^2}\left(\frac{6\pi^2}{2s+1}\frac{N}{V}\right)^{2/3}} - 1\right] \tag{5.79}$$

It may be shown, and the reader is invited to do so as an exercise, that from this expression we recover the non-relativistic expression in the limit $c \to \infty$. This relativistic case is useful in the treatment of white dwarf stars approaching the **Chandrasekhar mass** limit.

13. In fact, this definition would also apply in the classical regime since we then have $\epsilon_0 = 0$.

# 6 Ideal bosonic gas

In this chapter, we move to the treatment of an **ideal bosonic gas**, in which particles are identical bosons, with integer spin $s$, are independent, free from any external field, and non-relativistic.

## 6.1 General properties

### 6.1.1 The Bose factor

Similarly to what we did for the fermionic gas, we can treat the bosonic gas by converting the discrete sums over individual states $|\lambda\rangle$ to integrals over the energy $\epsilon$, using the density of states $\rho(\epsilon)$. Since the **average occupation number** for bosons is given by the Bose-Einstein distribution

$$\langle N_\lambda \rangle^B = \frac{1}{e^{\beta(\epsilon_\lambda - \mu)} - 1} \tag{6.1}$$

this prompts the introduction of the **Bose factor**

$$N^B(\epsilon; T, \mu) = \frac{1}{e^{\beta(\epsilon-\mu)} - 1} \tag{6.2}$$

In the following, we may omit the dependence on $T$ and $\mu$, and consider it implicitly. The shape of the function is shown in Fig. 6.1 for two different chemical potentials $\mu_1$ and $\mu_2$ with fixed temperature $T$, and for three different temperatures at fixed $\mu$. The Bose factor is a decreasing function of the energy, going to $+\infty$ when $\epsilon \to \mu$, and to zero when $\epsilon \to \infty$. We note that the expression above implies that the Bose factor is defined only for $\epsilon > \mu$, as already mentioned in (4.62).

### 6.1.2 Physical quantities

As we did in the case of the fermionic gas, we can work out the average number of particles in the grand-canonical framework (or the chemical potential in the canonical framework) through the integral

$$\langle N \rangle = \int_{\epsilon_0}^{\infty} \rho(\epsilon) \frac{1}{e^{\beta(\epsilon-\mu)} - 1} \mathrm{d}\epsilon \tag{6.3}$$

and the average energy through a similar integral

$$\langle E \rangle = \int_{\epsilon_0}^{\infty} \rho(\epsilon) \frac{\epsilon}{e^{\beta(\epsilon-\mu)} - 1} \mathrm{d}\epsilon \tag{6.4}$$

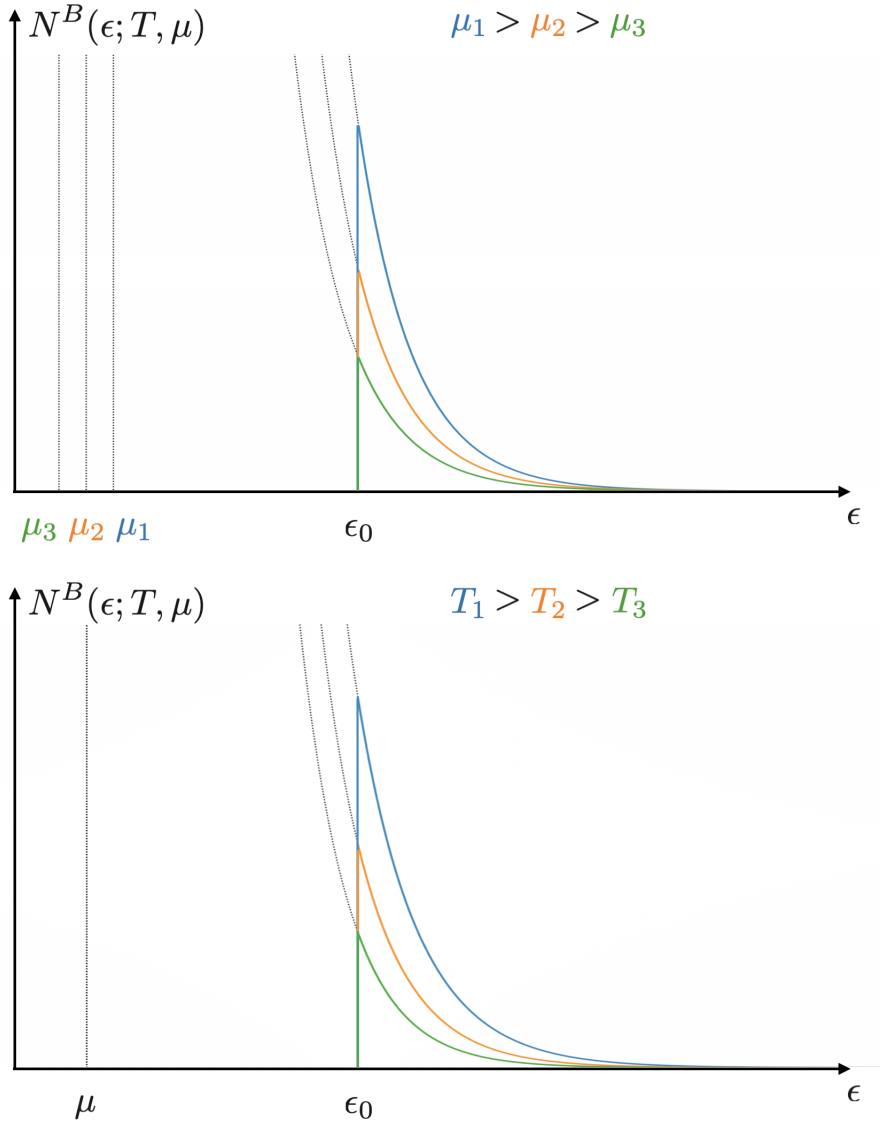



FIGURE 6.1 – The Bose factor, for different chemical potentials (top) and temperatures (bottom). The ground-level energy $\epsilon_0$ must be larger than the chemical potential.

The grand-canonical potential may be computed using a similar expression

$$J = -k_B T \int_{\epsilon_0}^{\infty} \rho(\epsilon) \ln \left[1 + N^B(\epsilon)\right] \mathrm{d}\epsilon \tag{6.5}$$

Indeed, we have the following relationship between the average occupation number and the individual state grand-canonical partition function in the bosonic case,

$$1 + \langle N_\lambda \rangle = 1 + \frac{1}{e^{\beta(\epsilon_\lambda - \mu)} - 1} = \frac{e^{\beta(\epsilon_\lambda - \mu)}}{e^{\beta(\epsilon_\lambda - \mu)} - 1} = \frac{1}{1 - e^{-\beta(\epsilon_\lambda - \mu)}} = \xi_\lambda \tag{6.6}$$

Other thermodynamic quantities may be computed in the same way. The reader is encouraged to try and compute the entropy $S$ as an exercise. We recall that

$$S = -\left(\frac{\partial J}{\partial T}\right)_{V,\mu} \tag{6.7}$$

### 6.1.3 Canonical interpretation

As already noted for the fermionic gas, the thermodynamic limit ensures that for large enough systems, the grand-canonical and canonical descriptions are equivalent, due to the very small range of fluctuations possible for internal variables, including $N$. Consequently, if we work in the grand-canonical ensemble, fixing $T$, $V$, and $\mu < \epsilon_0$, we may compute the average number $\langle N \rangle(T, V, \mu)$ as we did just then, and the relationship between these quantities may be (at least theoretically) inverted to yield $\mu(T, V, N)$ which would be the chemical potential computed for this system in the canonical ensemble. The difference with the fermionic case is that, for this to be correct, this solution $\mu(T, V, N)$ should verify the condition $\mu < \epsilon_0$, otherwise it would not be a physically sound solution to the problem in the bosonic case.

### 6.1.4 Example of the three-dimensional ideal gas

#### The Bose temperature

Specifiying the density of states $\rho(\epsilon)$ and the ground energy $\epsilon_0$ to be those of a non-relativistic ideal gas, we have the average number of particles in the grand-canonical framework given by

$$\langle N \rangle = \int_0^\infty \frac{AV\epsilon^{1/2}}{e^{\beta(\epsilon-\mu)} - 1} \mathrm{d}\epsilon \tag{6.8}$$

where we recall that

$$A = \frac{(2s+1)m^{3/2}}{\sqrt{2}\pi^2\hbar^3} \tag{6.9}$$

Introducing the fugacity $\varphi = e^{\beta\mu}$ and changing to the variable $x = \beta\epsilon$, we have

$$\frac{\langle N \rangle}{AV\left(k_B T\right)^{3/2}} = \int_0^\infty \frac{x^{1/2}}{\frac{e^x}{\varphi} - 1} \mathrm{d}x = I_B(\varphi) \tag{6.10}$$

The function $I_B(\varphi)$ is represented in Fig. 6.2. It is monotonously increasing from $I_B(0) = 0$ to a **finite value** at $\varphi = 1$. This is a fundamental property, because the condition that $\mu < \epsilon_0 = 0$ implies $\varphi < 1$. This finite value is [1]

$$I_B(1) = \Gamma\left(\frac{3}{2}\right)\zeta\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2} \times 2.612\ldots \approx 2.315\ldots \tag{6.14}$$

1. This integral makes use of Euler's Gamma function

$$\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}\mathrm{d}x \tag{6.11}$$

and Riemann's zeta function

$$\zeta(s) = \sum_{n \geqslant 1} \frac{1}{n^s} \tag{6.12}$$

for which we give a few important values

$$\Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2} \qquad \Gamma\left(\frac{5}{2}\right) = \frac{3\sqrt{\pi}}{4} \qquad \zeta\left(\frac{3}{2}\right) \approx 2.612\ldots \qquad \zeta\left(\frac{5}{2}\right) \approx 1.341\ldots \tag{6.13}$$

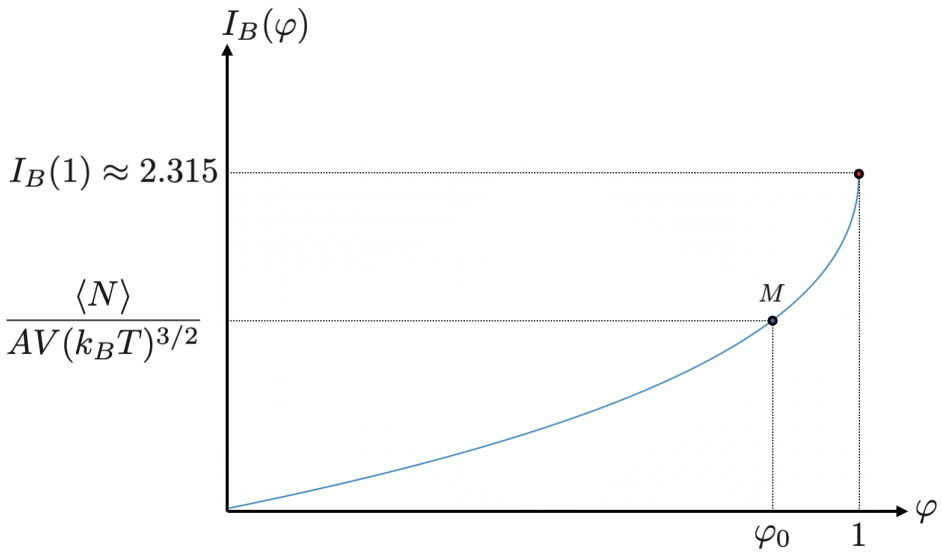



FIGURE 6.2 – The function $I_B(\varphi)$ and finding the solution $\varphi_0$ to the implicit Eq. (6.10).

Consequently, if we view the above equation as an implicit one on $\mu$ (or $\varphi$) in the canonical framework where the number of particles is fixed

$$\frac{N}{AV\left(k_BT\right)^{3/2}} = I_B(\varphi) \tag{6.15}$$

there cannot be a physical solution to the problem if $N > AV\left(k_BT\right)^{3/2} I_B(1)$. The admissibility condition may be written in terms of the temperature [2]

$$T < T_B \tag{6.16}$$

where we introduce the critical temperature known as the **Bose temperature**

$$T_B = \frac{2\pi\hbar^2}{mk_B}\left(\frac{1}{2s+1}\frac{1}{2.612\ldots}\frac{N}{V}\right)^{2/3} \tag{6.17}$$

We note that something must happen at this critical temperature, unlike in the case of the Fermi temperature $T_F$, that is just a scale with which to compare $T$ in order to know whether a classical or quantum treatment is in order, but where nothing fundamental happens.

### Properties at $T > T_B$

In that case, the treatment is very similar to the fermionic case, since there is a single $\varphi < 1$ satisfying the condition $N = AV\left(k_BT\right)^{3/2} I_B(\varphi)$. This yields the appropriate chemical potential $\mu$, and from there we may (at least theoretically and numerically) compute various thermodynamic quantities, starting with the energy

$$E = \int_0^\infty \frac{AV\epsilon^{3/2}}{e^{\beta(\epsilon-\mu)}-1}\mathrm{d}\epsilon \tag{6.18}$$

2. For a given $N$ and $V$, the representative point $M$ on the figure moves up the curve as $T$ decreases, and goes beyond $\varphi = 1$ for the critical temperature $T_B$.

The grand-canonical potential is computed as already noted

$$J = -k_B T \int_{\epsilon_0}^{\infty} \rho(\epsilon) \ln\left[1 + N^B(\epsilon)\right] \mathrm{d}\epsilon = k_B T A V \int_0^{\infty} \epsilon^{1/2} \ln\left[1 - e^{-\beta(\epsilon-\mu)}\right] \mathrm{d}\epsilon \tag{6.19}$$

An integration by parts of this last expression (this may be done as an exercise) yields

$$J = -\frac{2}{3} A V \int_0^{\infty} \frac{\epsilon^{3/2}}{e^{\beta(\epsilon-\mu)} - 1} \mathrm{d}\epsilon = -\frac{2}{3} E \tag{6.20}$$

Since $J = -PV$, a result derived from the linearity of $J$ with respect to $V$, we conclude that for any monoatomic, non-relativistic ideal gas, whether classical, fermionic, or bosonic, we have

$$PV = \frac{2}{3} E \tag{6.21}$$

The equation of state of the bosonic, non-relativistic ideal gas at $T > T_B$ is then known in a parametric way, where the parameter is the chemical potential $\mu$

$$P = \frac{(2s+1)\sqrt{2} m^{3/2}}{3\pi^2 \hbar^3} \int_0^{\infty} \frac{\epsilon^{3/2}}{e^{\beta(\epsilon-\mu)} - 1} \mathrm{d}\epsilon \qquad N = \frac{(2s+1) V m^{3/2}}{\sqrt{2}\pi^2 \hbar^3} \int_0^{\infty} \frac{\epsilon^{1/2}}{e^{\beta(\epsilon-\mu)} - 1} \mathrm{d}\epsilon \tag{6.22}$$

Indeed, we then have [3]

$$\frac{PV}{N} = \frac{\dfrac{(2s+1)\sqrt{2} m^{3/2}}{3\pi^2 \hbar^3} V \displaystyle\int_0^{\infty} \frac{\epsilon^{3/2}}{e^{\beta(\epsilon-\mu)} - 1} \mathrm{d}\epsilon}{\dfrac{(2s+1) m^{3/2}}{\sqrt{2}\pi^2 \hbar^3} V \displaystyle\int_0^{\infty} \frac{\epsilon^{1/2}}{e^{\beta(\epsilon-\mu)} - 1} \mathrm{d}\epsilon} = \frac{2}{3} \frac{\displaystyle\int_0^{\infty} \frac{\epsilon^{3/2}}{e^{\beta(\epsilon-\mu)} - 1} \mathrm{d}\epsilon}{\displaystyle\int_0^{\infty} \frac{\epsilon^{1/2}}{e^{\beta(\epsilon-\mu)} - 1} \mathrm{d}\epsilon} = f_\mu(T) \tag{6.23}$$

This can be written a bit more explicity as the following equation of state

$$PV = \frac{2}{3} \frac{H_B(\varphi)}{I_B(\varphi)} N k_B T \tag{6.24}$$

introducing the function [4]

$$H_B(\varphi) = \int_0^{\infty} \frac{x^{3/2}}{\dfrac{e^x}{\varphi} - 1} \mathrm{d}x \tag{6.26}$$

Other thermodynamic quantities may be derived from the grand-canonical potential, such as the entropy. The reader is invited to try and derive these in the form of integrals over the energy $\epsilon$.

## 6.2 Low-temperature behaviour

### 6.2.1 Bose condensation

As we mentioned, $T_B$ plays a particular role, and some physical process must occur at this temperature. This is unlike $T_F$, that only provides a scale for comparing $T$ to. What happens physically

3. Note that $\mu$ is a function of $T$, $V$, and $N$, so that the form here is indeed the equation of state $\mathcal{F}(P, T, V, N) = 0$.

4. Note that in the classical limit of high temperature, the Bose factor reduces to the Boltzmann one and we have

$$\frac{H_B(\varphi)}{I_B(\varphi)} = \frac{\displaystyle\int_0^{\infty} x^{3/2} e^{-x} \mathrm{d}x}{\displaystyle\int_0^{\infty} x^{1/2} e^{-x} \mathrm{d}x} = \frac{\Gamma\left(\frac{5}{2}\right)}{\Gamma\left(\frac{3}{2}\right)} = \frac{\frac{3}{4}\sqrt{\pi}}{\frac{1}{2}\sqrt{\pi}} = \frac{3}{2} \tag{6.25}$$

and we recover the classical equation of state $PV = N k_B T$.

may be understood relatively easily. As the temperature goes down, the particles tend to occupy lower energy states. While in the case of fermions, the Pauli exclusion principle forbade them to occupy the same state, and in particular the ground level, simultaneously, no such impossibility occurs for bosons. They may well occupy the ground state together. So much so that the number of bosons in that state may become macroscopic. This is the phenomenon known as **Bose condensation**. Yet, the density of states for $\epsilon = \epsilon_0 = 0$ is $\rho(0) = 0$ so these particles are simply forgotten when converting from the discrete sum to the integral over the energy

$$\langle N \rangle = \sum_{|\lambda\rangle} \langle N_\lambda \rangle^B \neq \int_0^\infty \rho(\epsilon) N^B(\epsilon) \mathrm{d}\epsilon \tag{6.27}$$

At low temperatures, the ground state is going to be much more populated than any excited state, so it should be treated separately. Indeed, its population is

$$N_0 = \frac{1}{e^{-\beta\mu} - 1} \approx -\frac{k_B T}{\mu} \gg 1 \tag{6.28}$$

which may be obtained from noticing that the exponential factor is $e^{-\beta\mu} \approx 1 - \beta\mu$ (we recall that $\mu \lesssim 0$ since the point $M$ on the curve of Fig. 6.2 approaches $\varphi \to 1^-$). The first excited state has an energy

$$\epsilon_1 = \frac{\hbar^2 k_1^2}{2m} \tag{6.29}$$

where $k_1$ is the wavenumber of a wavevector whose tip is at the innermost non-zero grid point of the lattice in $\vec{k}$ space. Assuming that the box is cubic, $V = L^3$, we have [5]

$$k_1 = \frac{2\pi}{L} \qquad \epsilon_1 = \frac{2\pi^2\hbar^2}{mL^2} \ll k_B T \tag{6.30}$$

The population of this first excited state is then, using the expression of the occupation number,

$$N_1 = \frac{1}{e^{\beta(\epsilon_1 - \mu)} - 1} \approx \frac{1}{\frac{\epsilon_1}{k_B T} - \frac{\mu}{k_B T}} \approx \frac{1}{\frac{\epsilon_1}{k_B T} + \frac{1}{N_0}} \tag{6.31}$$

since [6] $\beta(\epsilon_1 - \mu) \approx 0$. In the vicinity of $T_B$, the first ratio in the denominator is of the order

$$\frac{\epsilon_1}{k_B T} \approx \frac{\epsilon_1}{k_B T_B} = \frac{2\pi^2\hbar^2}{mL^2} \times \frac{m}{2\pi\hbar^2} \left[ (2s+1) \times 2.612 \times \frac{V}{N} \right]^{2/3} \approx \frac{C}{N^{2/3}} \tag{6.32}$$

where $C$ is a constant of order unity (in fact $C \approx 4$ for $s = 1$). This justifies that $\epsilon_1 \ll k_B T$ for a macroscopic system $N \gg 1$. Since we expect $N_0$ to be of order $N$, we have then

$$\frac{\epsilon_1}{k_B T} \approx \frac{C}{N^{2/3}} \gg \frac{1}{N_0} \tag{6.33}$$

and therefore the population of the first excited state is much smaller than that of the ground state,

$$N_1 \approx \frac{k_B T}{\epsilon_1} \approx \frac{N^{2/3}}{C} \ll N_0 \tag{6.34}$$

The Bose condensation only occurs in the ground state. Consequently, we can treat the problem as we did before, by converting the discrete sum over the individual states to an integral over $\epsilon$, provided that we treat the ground state separately,

5. The fact that $\epsilon_1 \ll k_B T$ is justified a bit further down.
6. We have already seen that $|\beta\mu| \ll 1$, and the first excited state has an energy $\epsilon_1 \ll k_B T$.

$$N = N_0 + \int_0^\infty \frac{AV\epsilon^{1/2}}{e^{\beta(\epsilon-\mu)}-1}\mathrm{d}\epsilon \tag{6.35}$$

The impression that this is insufficiently determined since there are apparently two unknowns ($N_0$ and $\mu$) is only superficial. Indeed $N_0$ is known, because the chemical potential must satisfy [7]

$$\mu = -\frac{k_B T}{N_0} \tag{6.36}$$

It is negative $\mu < 0$ and very small (compared to $k_B T$) in absolute value. In the integral, it is then readily approximated by $\mu = 0$, so that

$$N = N_0 + \int_0^\infty \frac{AV\epsilon^{1/2}}{e^{\beta\epsilon}-1}\mathrm{d}\epsilon = N_0 + AV\left(k_B T\right)^{3/2}\int_0^\infty \frac{x^{1/2}}{e^x-1}\mathrm{d}x \tag{6.37}$$

In the last integral we recognize, unsurprisingly, $I_B(1) \approx 2.315\ldots$, so

$$N = N_0 + AV\left(k_B T\right)^{3/2} I_B(1) \tag{6.38}$$

and inserting the expression of the Bose temperature, which is a function of $N$ and $V$ [8], we get the population of the ground state as a function of the total number of particles and of the temperature

$$N_0 = N\left[1 - \left(\frac{T}{T_B}\right)^{3/2}\right] \tag{6.39}$$

At $T = T_B$, this drops to zero, because the ground-state population ceases to be macroscopic. This is shown in Fig. 6.3.

### 6.2.2 Physical quantities at $T < T_B$

**Energy**

Since the ground-level energy is $\epsilon_0 = 0$, the energy of the system below the Bose temperature reads

$$E = N_0\epsilon_0 + \int_0^\infty \frac{AV\epsilon^{3/2}}{e^{\beta\epsilon}-1}\mathrm{d}\epsilon = AV\left(k_B T\right)^{5/2}\int_0^\infty \frac{x^{3/2}}{e^x-1}\mathrm{d}x \tag{6.40}$$

The integral is given in tables and we have

$$E = H_B(1)AV\left(k_B T\right)^{5/2} = \Gamma\left(\frac{5}{2}\right)\zeta\left(\frac{5}{2}\right)AV\left(k_B T\right)^{5/2} = \frac{3\sqrt{\pi}}{4}\times 1.341\times AV\left(k_B T\right)^{5/2} \tag{6.41}$$

in which we can replace the factor $AV$ using the fact that

$$N - N_0 = AV\left(k_B T\right)^{3/2} I_B(1) = AV\left(k_B T\right)^{3/2}\Gamma\left(\frac{3}{2}\right)\zeta\left(\frac{3}{2}\right) = N\left(\frac{T}{T_B}\right)^{3/2} \tag{6.42}$$

and therefore the energy $E$ of the bosonic ideal gas at $T < T_B$ is

$$E = \frac{\Gamma\left(\frac{5}{2}\right)\zeta\left(\frac{5}{2}\right)}{\Gamma\left(\frac{3}{2}\right)\zeta\left(\frac{3}{2}\right)}Nk_B T\left(\frac{T}{T_B}\right)^{3/2} \approx 0.77Nk_B T\left(\frac{T}{T_B}\right)^{3/2} \tag{6.43}$$

7. In other words, we have two equations for two unknowns.
8. We recall that $N = AV\left(k_B T_B\right)^{3/2} I_B(1)$.

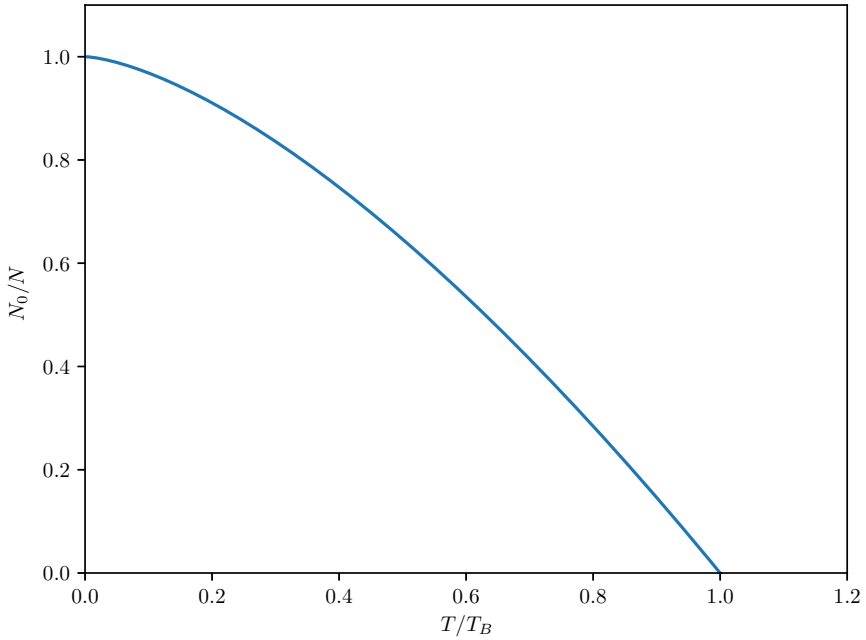



FIGURE 6.3 – Ground-level population $N_0$ as a function of $T$

At the transition $T = T_B$, the energy is therefore of the same order as that of the classical ideal gas $\sim Nk_BT_B$, but for $T < T_B$, the energy is equivalent to that of the classical ideal gas in which the ground-state particles would be removed, $E \sim (N - N_0)k_BT$. This is not surprising since the condensed particles cease to participate in the energy of the system ($\epsilon_0 = 0$).

### Heat capacity

The heat capacity results from the derivation of the above expression with respect to temperature,

$$C_v \approx 1.925 N k_B \left(\frac{T}{T_B}\right)^{3/2} \tag{6.44}$$

and is therefore of the same order as that of the classical ideal gas at the transition, $C_v(T_B) \sim Nk_B$, but is proportional to the non-condensed part of the gas below the transition, $C_v \sim (N - N_0)k_B$.

### Grand-canonical potential

The grand-canonical potential is also given by a sum over individual states that must be treated as the sum of the ground state term and an integral over the energy for the excited states

$$J = -k_BT \sum_{|\lambda\rangle} \ln \xi_\lambda = -k_BT \ln(1 + N_0) + k_BTAV \int_0^\infty \epsilon^{1/2} \ln\left[1 - e^{-\beta\epsilon}\right] \mathrm{d}\epsilon \tag{6.45}$$

Obviously, the first term may be approximated by $-k_B T \ln N_0$ so it is of order a few tens of $k_B T$, and therefore completely negligible with respect to the second term, that is of order $N k_B T$. Thus,

$$J = k_B T A V \int_0^\infty \epsilon^{1/2} \ln \left[1 - e^{-\beta\epsilon}\right] \mathrm{d}\epsilon \tag{6.46}$$

An integration by parts then yields the same relation as was obtained for $T > T_B$,

$$J = -\frac{2}{3} E \tag{6.47}$$

**Helmholtz free energy**

The Helmholtz free energy may be computed from the grand-canonical potential as $F = J + \mu N$. Now in the domain where Bose condensation occurs, the chemical potential is null [9], so

$$F = -\frac{2}{3} E \tag{6.48}$$

**Entropy**

The entropy is obtained from the Helmholtz free energy [10]

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} = \frac{2}{3} C_v \propto (N - N_0) \tag{6.49}$$

showing that the $N_0$ condensed particles do not provide any entropy to the system, which is not surprising, as their state is perfectly known.

**Pressure**

The pressure is obtained from the relation stemming from the linearity of $J$ with respect to $V$,

$$P = -\frac{J}{V} = \frac{2}{3}\frac{E}{V} \approx 1.341(2s+1)\left(\frac{m}{2\pi\hbar^2}\right)^{3/2} (k_B T)^{5/2} \tag{6.50}$$

This is quite astonishing that the pressure only depends on the temperature and not on the density $N/V$. The equation of state for an ideal bosonic gas below the Bose condensation point is a relation between $P$, and $T$, irrespective of the density.

9. Since in fact $\mu = -(k_B T)/N_0$, the second term is of order $k_B T$, much smaller than $J$.
10. There is no such problem as in the fermionic case, because here the chemical potential is essentially "fixed" at zero below the Bose temperature.

# 7 Thermodynamics of radiation

In this chapter, we discuss the thermodynamics of a **photon gas**, and the associated notion of **black-body radiation**. This is of course of great importance in astrophysical contexts. As we will see, a photon gas can be treated as a special case of a bosonic gas in which the chemical potential is null, due to the photon number not being a conserved quantity.

## 7.1 Photon statistics

### 7.1.1 Electromagnetic modes as photons

Consider a rectangular box of volume $V = L_x L_y L_z$, in which electromagnetic waves are present. Due to the linearity of Maxwell's equations, the $\vec{\mathcal{E}}$ and $\vec{\mathcal{B}}$ fields within the box may be written as a superposition of monochromatic plane waves of the form

$$\vec{\mathcal{E}}(\vec{r},t) \propto \vec{\tilde{\mathcal{E}}}\left(\vec{k},\omega\right) e^{i\left(\vec{k}.\vec{r}-\omega t\right)} \qquad \vec{\mathcal{B}}(\vec{r},t) \propto \vec{\tilde{\mathcal{B}}}\left(\vec{k},\omega\right) e^{i\left(\vec{k}.\vec{r}-\omega t\right)} \tag{7.1}$$

The periodic conditions on the boundaries of the box, induced by the fact that multiple reflections on them lead to standing waves, impose that the wavevector must be quantized

$$\vec{k} = \frac{2\pi n_x}{L_x}\vec{e}_x + \frac{2\pi n_y}{L_y}\vec{e}_y + \frac{2\pi n_z}{L_z}\vec{e}_z \tag{7.2}$$

with $(n_x, n_y, n_z) \in \mathbb{Z}^3$. This wavevector defines a **mode** of the electromagnetic field. From Maxwell's equations, it may be shown that $\vec{\mathcal{E}}$ and $\vec{\mathcal{B}}$ obey partial differential equations such as

$$\Delta\vec{\mathcal{E}} - \frac{1}{c^2}\frac{\partial^2\vec{\mathcal{E}}}{\partial t^2} = 0 \tag{7.3}$$

for which non-trivial solutions exist only if $\vec{k}$ and $\omega$ are related, through the **dispersion relation**. Consequently, the amplitudes of the modes can be written as functions of the sole wavevector, $\vec{\tilde{\mathcal{E}}}\left(\vec{k}\right)$ and $\vec{\tilde{\mathcal{B}}}\left(\vec{k}\right)$. Now, in classical physics (i.e., classical electromagnetism), the energy contained in a mode is proportional to its squared amplitude, so the energy carried by the radiation in the box is [1]

$$E \propto \sum_{\vec{k}} \left|\vec{\tilde{\mathcal{E}}}\left(\vec{k}\right)\right|^2 \tag{7.4}$$

1. Note that more rigorously one should consider the squared amplitude of the full electric field, as a superposition of the modes. The cross-terms that appear for different modes $\vec{k}$ and $\vec{k}'$ in fact yield a null contribution, leading to Eq. (7.4).

The contribution to this sum from any mode $\vec{k}$ can take any (continuous) value from 0 to $\infty$, and this leads to major problems that can only be solved by taking a quantum approach in which the energy in each mode is quantized. This is the revolutionary idea introduced by Planck : to each monochromatic electromagnetic plane wave we should associate a particle called **photon** that has an energy $\epsilon$ and a momentum $\vec{p}$ related to the properties $\omega$ and $\vec{k}$ of the wave by

$$\epsilon = \hbar\omega \qquad \vec{p} = \hbar\vec{k} \tag{7.5}$$

and that in each mode there can only be an integer ($n = 0, 1, 2, \ldots$) number of photons, so that the energy associated to that mode is not any value but $n\hbar\omega$. In vacuum, the dispersion relation we mentioned above is

$$\omega = c\left|\vec{k}\right| = ck \tag{7.6}$$

where $c = 299\,792\,458\,\mathrm{m\,s^{-1}}$ is the speed of light. The relation between the energy and the momentum of photons is then

$$\epsilon = c\,|\vec{p}| = \hbar ck \tag{7.7}$$

which is a special case of the relativistic relationship in the case of massless particles, such as photons

$$\epsilon = \sqrt{p^2c^2 + m^2c^4} \tag{7.8}$$

To the two independent polarizations possible for the electromagnetic wave (e.g., right- and left-handed circular polarization) correspond two **spin states** for the photon. However, this does not mean that $s = 1/2$. Instead the spin of the photon is unity, $s = 1$ but the possible values of the spin projection on a given axis are $s_z = \pm 1$, excluding the value 0. This is a general property for massless particles. Consequently, **photons are bosons**.

### 7.1.2 Density of states

The quantization of the wavevector and the linear relationship between wavenumber and energy of the photon, $\epsilon = \hbar ck$, allow to determine easily the cumulative density of states of the photon in the box, using the same approach as in (5.3.1). We have, taking into account the spin states,

$$\Phi(\epsilon) = 2 \times \frac{\frac{4}{3}\pi k(\epsilon)^3}{\frac{2\pi}{L_x} \times \frac{2\pi}{L_y} \times \frac{2\pi}{L_z}} = \frac{Vk(\epsilon)^3}{3\pi^2} = \frac{V\epsilon^3}{3\pi^2(\hbar c)^3} \tag{7.9}$$

and deriving this with respect to $\epsilon$ we have the density of states

$$\rho(\epsilon) = \frac{V\epsilon^2}{\pi^2(\hbar c)^3} \tag{7.10}$$

One should note the difference with the case of an ideal gas of material particles, either relativistic or not. For instance, we recall that the densities of states for a non-relativistic gas are $\Phi(\epsilon) \propto \epsilon^{3/2}$ and $\rho(\epsilon) \propto \epsilon^{1/2}$.

### 7.1.3 Distribution function of photons

Consider that the box containing the radiation is in thermal equilibrium at temperature $T$, by being in contact with a thermostat $\mathcal{T}$ imposing that temperature. The unit spin means that photons are

bosons, and the linearity of Maxwell's equations implies that there is no interaction between photons, so the system may be viewed as an ideal bosonic gas. Also, the number of photons is not conserved as these are constantly being absorbed and emitted by the walls of the box, in contrast with the treatment of the boson gas in Chap. 6.

What we are describing here is a **black body** : a closed cavity containing an ensemble of photons in equilibrium with the walls, maintained at a given temperature $T$. For the photonic gas to be considered ideal, the coupling between it and the cavity must be weak, but it must exist, because the equilibrium between photons cannot be attained in the absence of a mediating body since they do not interact. They achieve equilibrium among themselves by interacting with the box, i.e., being absorbed and emitted by the walls. The non-conservation of the number $N$ of photons, contrary to what occurs in all gases we have considered so far, implies that this quantity should be determined based on equilibrium conditions. We saw that in the context of imposing volume and temperature (the canonical ensemble [2]), a necessary condition for equilibrium was for the Helmholtz free-energy $F$ to be minimal. It should therefore be minimal with respect to $N$, which is in this case an internal variable. Because of this, we see from equation (3.39) that the chemical potential is null

$$\mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} = 0 \tag{7.11}$$

We may therefore conclude that the photon gas in thermal equilibrium is described by the Bose-Einstein distribution for $\mu = 0$,

$$\langle N_\lambda \rangle = \frac{1}{e^{\beta\epsilon_\lambda} - 1} \tag{7.12}$$

where $|\lambda\rangle$ stands for the individual states of photons (wavevector $\vec{k}$ and spin, which can be either "left" or "right", for instance). We note that there cannot be any Bose condensation here since this requires particle number conservation [3].

## 7.2 Laws of thermal radiation

### 7.2.1 Planck's law

#### Determination of the law

Using the density of states $\rho(\epsilon)$ of equation (7.10) and the average occupation number just determined, we can obtain the number of photons in the box whose energy lies within $[\epsilon, \epsilon + \mathrm{d}\epsilon]$ as

$$\mathrm{d}N(\epsilon, T) = \frac{\rho(\epsilon)\mathrm{d}\epsilon}{e^{\beta\epsilon} - 1} = \frac{V}{\pi^2(\hbar c)^3} \frac{\epsilon^2 \mathrm{d}\epsilon}{e^{\beta\epsilon} - 1} \tag{7.13}$$

These photons carry an energy that is simply that number times the energy of a photon

$$\mathrm{d}E(\epsilon, T) = \epsilon \mathrm{d}N(\epsilon, T) = \frac{V}{\pi^2(\hbar c)^3} \frac{\epsilon^3 \mathrm{d}\epsilon}{e^{\beta\epsilon} - 1} \tag{7.14}$$

From this expression, we may derive the **spectral energy density** in pulsation, per unit volume $u_\omega(T)$ such that $\mathrm{d}E(\epsilon, T) = V u_\omega(T)\mathrm{d}\omega$. Quite easily, we find that this quantity obeys the **Planck law**

2. Strictly speaking, in the canonical ensemble all three $T$, $V$, and $N$ are fixed external parameters. Here, only $T$ and $V$ are such, and $N$ is actually an internal variable, which justifies the approach taken here to prove that $\mu = 0$.

3. Actually, there have been observations of Bose condensations for photons in very specific settings by which photon conservation was obtained as a prerequisite (see, e.g., Klaers et al., Nature, 468, 545, 2010).

$$u_\omega(T) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{\exp\left(\frac{\hbar\omega}{k_B T}\right) - 1} \tag{7.15}$$

### Other expressions of the Planck law

There are many other, equivalent ways to express this result. For instance, this spectral energy distribution may be expressed in frequency $\nu$ rather than in pulsation, using the relation $\nu = \omega/(2\pi)$ and the condition that $u_\nu(T)\mathrm{d}\nu = u_\omega(T)\mathrm{d}\omega$ for equivalent ranges

$$u_\nu(T) = \frac{8\pi h\nu^3}{c^3} \frac{1}{\exp\left(\frac{h\nu}{k_B T}\right) - 1} \tag{7.16}$$

This may also be expressed per unit wavelength $\lambda = c/\nu$ instead of per unit frequency. Using the constraint $u_\nu|\mathrm{d}\nu| = u_\lambda|\mathrm{d}\lambda|$, where the absolute values ensure both $u_\nu$ and $u_\lambda$ to be positive, we have

$$u_\lambda(T) = \frac{8\pi hc}{\lambda^5} \frac{1}{\exp\left(\frac{hc}{\lambda k_B T}\right) - 1} \tag{7.17}$$

Another quite common expression of the same result is to write not the energy per unit volume and unit frequency interval $u_\nu$ or unit wavelength interval $u_\lambda$, but the **specific intensity** or **brightness** $I_\nu$ (or $I_\lambda$), which is a fundamental quantity in astrophysics and is defined as the power per unit surface, per unit solid angle, and per unit frequency (or wavelength) interval. In the case of an isotropic radiation field such as that of a black body, we have

$$I_\nu = \frac{c}{4\pi} u_\nu \qquad\qquad I_\lambda = \frac{c}{4\pi} u_\lambda \tag{7.18}$$

and therefore the Planck law, written in terms of these, states that $I_\nu = B_\nu(T)$ and $I_\lambda = B_\lambda(T)$ with

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{\exp\left(\frac{h\nu}{k_B T}\right) - 1} \tag{7.19}$$

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5} \frac{1}{\exp\left(\frac{hc}{\lambda k_B T}\right) - 1} \tag{7.20}$$

### Representative curves

At a given frequency $\nu$, if $T_2 > T_1$, we immediately have $B_\nu(T_2) > B_\nu(T_1)$ and $B_\lambda(T_2) > B_\lambda(T_1)$, showing that the representative curves of black body emission are like "Russian dolls". These curves, for $B_\nu$, are shown for various temperatures in Fig. 7.1.

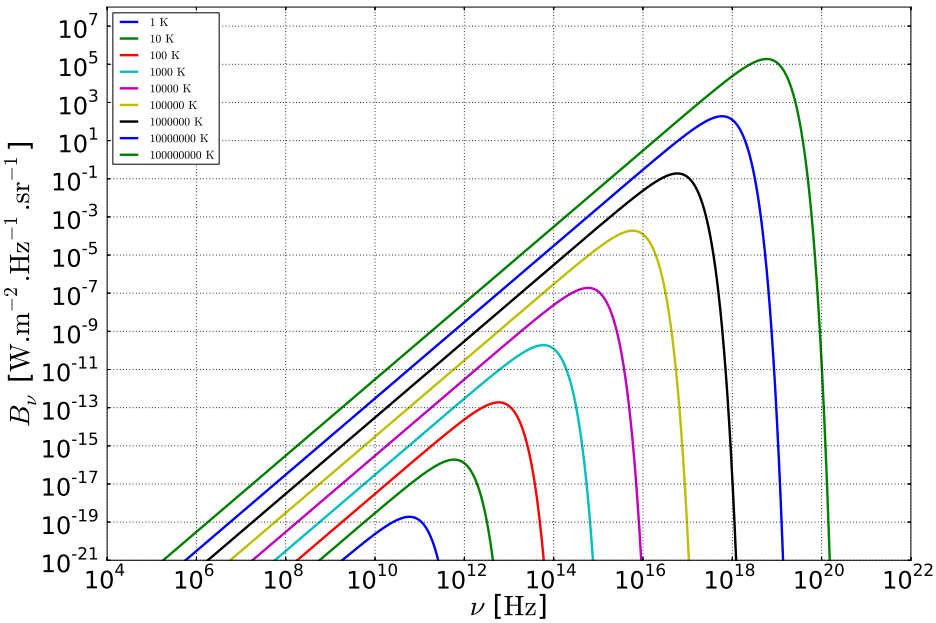



FIGURE 7.1 – Curves for black body emission $B_\nu$ at various temperatures, as a function of frequency.

### 7.2.2 Limiting cases

**Low frequencies : Rayleigh-Jeans law**

At low frequency, i.e. for $\nu \ll k_B T/h$, or equivalently long wavelengths $\lambda \gg hc/k_B T$, Planck functions $B_\nu(T)$ and $B_\lambda(T)$ may be approximated by the following expressions

$$B_{\nu,\mathrm{RJ}} = \frac{2k_B T\nu^2}{c^2} \qquad B_{\lambda,\mathrm{RJ}} = \frac{2k_B cT}{\lambda^4}. \tag{7.21}$$

These form the **Rayleigh-Jeans approximation**, that is particularly useful in radioastronomy. We note that the spectral energy density in pulsation associated to that approximation is

$$u_{\omega,\mathrm{RJ}} = \frac{k_B T\omega^2}{\pi^2 c^3}, \tag{7.22}$$

an expression that may be interpreted classically. Indeed, recall that in classical physics, energy equipartition states that an energy $k_B T$ is ascribed to each 1-dimensional oscillator in thermal equilibrium [4]. Therefore, the energy per unit volume associated to modes of the radiation field within the range $[\omega, \omega + \mathrm{d}\omega]$ is, classically,

$$u_{\omega,\mathrm{cl}}\mathrm{d}\omega = \frac{\tilde{\rho}(\omega)\mathrm{d}\omega}{V} k_B T \tag{7.23}$$

where the density of modes $\tilde{\rho}(\omega)$ is obtained by equating $\tilde{\rho}(\omega)\mathrm{d}\omega$ to the corresponding number in energy space $\rho(\epsilon)\mathrm{d}\epsilon$, i.e.,

$$\tilde{\rho}(\omega) = \frac{V\epsilon^2}{\pi^2(\hbar c)^3}\frac{\mathrm{d}\epsilon}{\mathrm{d}\omega} = \frac{V\omega^2}{\pi^2 c^3} \tag{7.24}$$

4. We have $k_B T/2$ for the mean kinetic energy and $k_B T/2$ for the mean potential energy.

from which we have $u_{\omega,\mathrm{cl}} = u_{\omega,\mathrm{RJ}}$. The Rayleigh-Jeans form is what we would find classically, and we see that the total radiation energy in the box, integrated over all pulsations, becomes infinite

$$u_{\mathrm{cl}} = \int_0^\infty u_{\omega,\mathrm{cl}} \mathrm{d}\omega = \frac{k_B T}{\pi^2 c^3} \int_0^\infty \omega^2 \mathrm{d}\omega = \infty \tag{7.25}$$

This is the **ultraviolet catastrophe** that prompted the suggestion for a quantization of the radiation field. This quantization is not apparent at low frequencies $h\nu \ll k_B T$ because there needs to be a great many photons with energy $h\nu$ to attain the typical thermal energy $k_B T$ in a given mode.

#### High frequencies : Wien law

Conversely, at high frequencies $\nu \gg k_B T/h$ or short wavelengths $\lambda \ll hc/k_B T$, the Planck functions $B_\nu(T)$ and $B_\lambda(T)$ may be approximated by the following expressions,

$$B_{\nu,\mathrm{W}} = \frac{2h\nu^3}{c^2} \exp\left(-\frac{h\nu}{k_B T}\right) \qquad B_{\lambda,\mathrm{W}} = \frac{2hc^2}{\lambda^5} \exp\left(-\frac{hc}{\lambda k_B T}\right). \tag{7.26}$$

These form the **Wien approximation**. In this limit, even just one photon provides an energy much larger than $k_B T$, so there cannot be many photons in these modes at thermal equilibrium. The figure below shows the Planck law for a black-body at the temperature of the Sun, as well as the Rayleigh-Jeans and Wien approximations that are appropriate for low and high frequencies, respectively.

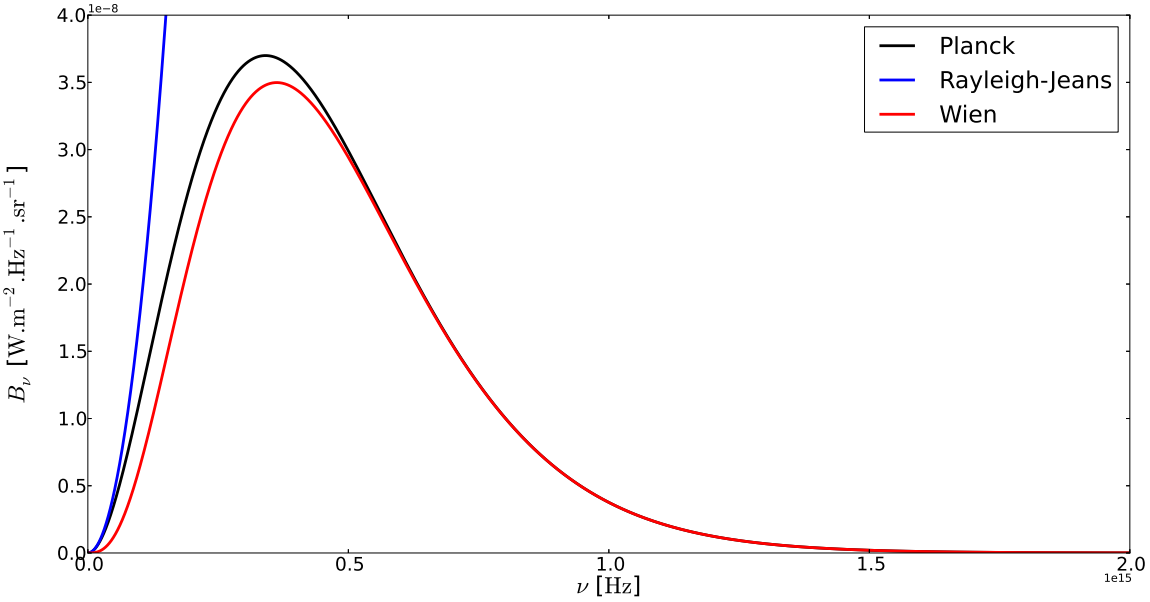



FIGURE 7.2 – Planck's, Rayleigh-Jeans' and Wien's laws for black body emission at $T = 5800\,\mathrm{K}$.

### 7.2.3 Wien's displacement law

The representative curves for black body emission present a maximum, that may be computed semi-analytically. We discuss here the case of $B_\nu(T)$, but similar results may be obtained for $B_\lambda$. By posing

$x = h\nu/k_BT$, Planck's law may be rewritten as

$$B_\nu = \frac{2h\nu^3}{c^2}\frac{1}{\exp\left(\dfrac{h\nu}{k_BT}\right)-1} = \frac{2h}{c^2}\left(\frac{k_BT}{h}\right)^3\frac{x^3}{e^x-1} = \frac{2k_B^3T^3}{c^2h^2}\frac{x^3}{e^x-1} \tag{7.27}$$

so that the maximum of $B_\nu$ is obtained by finding that of $x^3/(e^x-1)$, a rather simple job numerically. The result is $x_{\max} \approx 2.82$, so that the frequency $\nu_{\max}$ where the black body emission $B_\nu(T)$ peaks is proportional to $T$, according to **Wien's displacement law**

$$\frac{\nu_{\max}}{T} \approx 2.82\frac{k_B}{h} = 5.88\,10^{10}\ \mathrm{Hz.K^{-1}} \tag{7.28}$$

This may be done also for $B_\lambda(T)$, yielding that the wavelength for which this is maximum is inversely proportional to $T$,

$$\lambda_{\max}T = 2.89\ 10^{-3}\ \mathrm{K.m} \tag{7.29}$$

We note that $\lambda_{\max}\,\nu_{\max} \neq c$ because the functions to be maximized are different. We may also consider the case of the function[5] $\nu B_\nu = \lambda B_\lambda$, for which we obtain a peak frequency $\nu'_{\max}$ and wavelength $\lambda'_{\max}$ that are related by $\lambda'_{\max}\nu'_{\max} = c$. Numerically, we have

$$\frac{\nu'_{\max}}{T} \approx 3.92\frac{k}{h} = 8.17\,10^{10}\ \mathrm{Hz.K^{-1}} \qquad \lambda'_{\max}T = 3.67\ 10^{-3}\ \mathrm{K.m} \tag{7.30}$$

As an example, for stars of different spectral types, including the Sun (G2), we have the following values for the peak :

| Spectral type | O5 | B5 | A5 | F5 | G2 | G5 | K5 | M5 |
|---|---|---|---|---|---|---|---|---|
| $T$ [K] | 54000 | 15200 | 8310 | 6700 | 5778 | 5660 | 4400 | 3200 |
| $\nu_{\max}$ [$10^{15}$ Hz] | 3.17 | 0.89 | 0.49 | 0.39 | 0.34 | 0.33 | 0.26 | 0.19 |
| $\lambda_{\max}$ [$\mu$m] | 0.05 | 0.19 | 0.35 | 0.43 | 0.50 | 0.51 | 0.66 | 0.91 |
| $\nu'_{\max}$ [$10^{15}$ Hz] | 4.41 | 1.24 | 0.68 | 0.55 | 0.47 | 0.46 | 0.36 | 0.26 |
| $\lambda'_{\max}$ [$\mu$m] | 0.07 | 0.24 | 0.44 | 0.55 | 0.64 | 0.65 | 0.83 | 1.15 |
| Domain | EUV | Far-UV | Near-UV | Blue | Yellow | Yellow | Red | Near IR |

### 7.2.4 Stefan-Boltzmann's law

We may compute the total energy contained in the radiation field as the integral of $\mathrm{d}E$

$$E(T,V) = \frac{V}{\pi^2(\hbar c)^3}\int_0^\infty \frac{\epsilon^3\mathrm{d}\epsilon}{\exp\left(\dfrac{\epsilon}{k_BT}\right)-1} = \frac{V(k_BT)^4}{\pi^2(\hbar c)^3}\int_0^\infty \frac{x^3\mathrm{d}x}{e^x-1} \tag{7.31}$$

The integral in the last expression is

$$\int_0^\infty \frac{x^3\mathrm{d}x}{e^x-1} = \frac{\pi^4}{15} \tag{7.32}$$

5. The relationship stems from the constraint that the energy $B_\nu|\mathrm{d}\nu|$ in the range $|\mathrm{d}\nu|$ should be equal to that $B_\lambda|\mathrm{d}\lambda|$ in the corresponding range $|\mathrm{d}\lambda|$.

so that we finally have the energy $E(T,V)$ given by **Stefan-Boltzmann's law**

$$E(T,V) = \frac{4V\sigma_S}{c}T^4 \qquad \sigma_S = \frac{2\pi^5 k_B^4}{15c^2h^3} = 5.67\ 10^{-8}\ \mathrm{W.m^{-2}.K^{-4}} \tag{7.33}$$

where $\sigma_S$ is **Stefan's constant**. The energy of the radiation field per unit volume is

$$u = \frac{E}{V} = \frac{4\sigma_S}{c}T^4 \tag{7.34}$$

From the expression of the energy $E$ we find the heat capacity

$$C_v = \left(\frac{\partial E}{\partial T}\right)_V = \frac{16\sigma_S}{c}VT^3 \tag{7.35}$$

The reason for isolating the factor $4/c$ out of the constant $\sigma_S$ is because Stefan-Boltzmann's law is often expressed in terms of the **bolometric outward flux** $\mathcal{F}$ of the black body, that is the power radiated per unit surface of the black body, integrated over all frequencies

$$\mathcal{F} = \int_0^\infty \mathrm{d}\nu \iint B_\nu(T)\cos\theta \mathrm{d}\Omega \tag{7.36}$$

where the double integral is over a half-space $\theta \in [0, \pi/2]$ and $\phi \in [0, 2\pi]$. This quantity is then given by the following form of Stefan-Boltzmann's law

$$\mathcal{F} = \sigma_S T^4 \tag{7.37}$$

### 7.2.5 Grand-canonical potential, radiation pressure, and entropy

From the average occupation number we derive the grand-canonical potential for the photon gas [6]

$$J = -k_BT\int_0^\infty \rho(\epsilon)\ln\left[1 + \frac{1}{\exp\left(\frac{\epsilon}{k_BT}\right) - 1}\right]\mathrm{d}\epsilon = \frac{k_BTV}{\pi^2(\hbar c)^3}\int_0^\infty \epsilon^2 \ln\left[1 - e^{-\beta\epsilon}\right]\mathrm{d}\epsilon \tag{7.39}$$

and this may be integrated by parts to reveal the relationship with the energy. This is similar to the computation of $J$ in the case of the bosonic gas in the previous chapter, with the difference that in that case $\rho(\epsilon) \propto \epsilon^{1/2}$ while in the present case $\rho(\epsilon) \propto \epsilon^2$. We find that

$$J = -\frac{1}{3}\frac{V}{\pi^2(\hbar c)^3}\int_0^\infty \frac{\epsilon^3}{e^{\beta\epsilon} - 1}\mathrm{d}\epsilon = -\frac{1}{3}E = -\frac{4V\sigma_S}{3c}T^4 \tag{7.40}$$

Note that the Helmholtz free energy is identical to the grand-canonical potential, since the chemical potential is null,

$$F = J + \mu N = J \tag{7.41}$$

6. Recall that we have

$$J = -k_BT\ln\Xi = -k_BT\sum_{|\lambda\rangle}\ln\xi_\lambda = -k_BT\sum_{|\lambda\rangle}\ln\left(1 + \langle N_\lambda\rangle\right) \tag{7.38}$$

Since $J = -PV$, a result derived from the linearity of $J$ with respect to $V$, we conclude the following important result for the **radiation pressure** of an ideal photonic gas

$$P = \frac{1}{3}\frac{E}{V} = \frac{u}{3} = \frac{4\sigma_S}{3c}T^4 \tag{7.42}$$

In fact, the result $P = u/3$ is valid for any radiation field that is isotropic [7]. As for the entropy of the photonic gas, it reads [8]

$$S = -\left(\frac{\partial J}{\partial T}\right)_V = \frac{1}{3}\left(\frac{\partial E}{\partial T}\right)_V = \frac{16\sigma_S}{3c}VT^3 \tag{7.43}$$

For an adiabatic, reversible transformation of the photon gas, the product $T^3V$ is therefore a constant. And since $P \propto T^4$ we find that in such a transformation $PV^{4/3}$ is also constant. We find a polytropic behaviour with an exponent $\gamma = 4/3$ that is also what we find in the extreme-relativistic case for a degenerate Fermi gas, such as in white dwarfs reaching the Chandrasekhar limit.

### 7.2.6 Thermodynamic derivation

One can obtain Stefan-Boltzmann's law and the form of the entropy of the photon gas through arguments of purely classical thermodynamics. Let us consider a box maintained at temperature $T$, in which a photon gas is at thermal equilibrium with the box, the latter being closed on one end by a piston, so that the volume $V$ may vary depending on the pressure exerted by the photon gas. The internal energy of the gas is $U(T,V) = u(T)V$, where the energy density $u$ is an intensive quantity that is uniform and independent of $V$. In an infinitesimal, reversible transformation of the system, the first principle gives

$$\mathrm{d}U = \delta Q + \delta W = (C_v \mathrm{d}T + l\mathrm{d}V) - p\,\mathrm{d}V = C_v\mathrm{d}T + (l-p)\mathrm{d}V = C_v\mathrm{d}T + \left(l - \frac{u}{3}\right)\mathrm{d}V \tag{7.44}$$

where we have used that the pressure $p = u/3$ for an isotropic radiation field [9] On the other hand,

$$\mathrm{d}U = \mathrm{d}\left(uV\right) = V\mathrm{d}u + u\mathrm{d}V = V\frac{\mathrm{d}u}{\mathrm{d}T}\mathrm{d}T + u\mathrm{d}V \tag{7.46}$$

so that

$$C_v = V\frac{\mathrm{d}u}{\mathrm{d}T} \quad \text{et} \quad l = \frac{4u}{3} \tag{7.47}$$

Using then the second principle of thermodynamics, we have

$$\mathrm{d}S = \frac{\delta Q}{T} = \frac{C_v}{T}\mathrm{d}T + \frac{l}{T}\mathrm{d}V = \frac{V}{T}\frac{\mathrm{d}u}{\mathrm{d}T}\mathrm{d}T + \frac{4u}{3T}\mathrm{d}V \tag{7.48}$$

Maxwell's relations state that cross-derivatives must be equal for the exact differential $\mathrm{d}S$, so

$$\left[\frac{\partial}{\partial T}\left(\frac{4u}{3T}\right)\right]_V = \left[\frac{\partial}{\partial V}\left(\frac{V}{T}\frac{\mathrm{d}u}{\mathrm{d}T}\right)\right]_T \Rightarrow \frac{4}{3T^2}\left[T\frac{\mathrm{d}u}{\mathrm{d}T} - u\right] = \frac{1}{T}\frac{\mathrm{d}u}{\mathrm{d}T} \Rightarrow \frac{1}{T}\frac{\mathrm{d}u}{\mathrm{d}T} = \frac{4u}{T^2} \tag{7.49}$$

7. We note that the relation $PV = E/3$ is also found for an extreme relativistic electron gas, as it should since the relationship between energy and momentum is then the same, $\epsilon = pc$.

8. There is no issue here of whether $\mu$ or $N$ is held constant, as the energy depends only on temperature and volume.

9. This is demonstrated for the black-body in the previous section, but is actually valid for any isotropic radiation field, because of the general defintions of the energy density $u$ and of the radiation pressure $p$

$$u = \frac{1}{c}\int_0^\infty \mathrm{d}\nu \iint I_\nu \mathrm{d}\Omega \qquad p = \frac{1}{c}\int_0^\infty \mathrm{d}\nu \iint I_\nu \cos^2\theta \mathrm{d}\Omega \tag{7.45}$$

where the double integrals are over all $4\pi$ steradian. The isotropy allows to take out the specific intensity $I_\nu$ from that double integral and the result ensues from the numerical values of the integrals over $\theta$.

which yields that

$$\frac{\mathrm{d}u}{u} = 4\frac{\mathrm{d}T}{T} \Rightarrow u = aT^4 \tag{7.50}$$

This in turns gives the expression of the entropy of the photon gas

$$\mathrm{d}S = \frac{V}{T}\frac{\mathrm{d}u}{\mathrm{d}T}\mathrm{d}T + \frac{4u}{3T}\mathrm{d}V = 4aVT^2\mathrm{d}T + \frac{4aT^3}{3}\mathrm{d}V = \frac{4a}{3}\left(3VT^2\mathrm{d}T + T^3\mathrm{d}V\right) = \frac{4a}{3}\mathrm{d}\left(VT^3\right) \tag{7.51}$$

i.e., using the third principle of thermodynamics

$$S = \frac{4a}{3}VT^3 \tag{7.52}$$

## 7.3 Emission and absorption of radiation

As we mentioned, the establishment of an equilibrium within the photon gas requires that it be mediated by interactions with matter. Material particles (atoms and molecules) harbour discrete energy levels (electronic, rotational, vibrational), and such a particle may undergo a transition from one level to the other by absorbing or emitting a photon of the right energy, that is, equal to the energy difference between those two levels. The probabilities for such transitions are ruled by a set of coefficients, called **Einstein coefficients**, that may be derived from one another through simple relations.

### 7.3.1 Absorption, spontaneous emission, stimulated emission

As a simple case, consider a two-level system, where the levels are noted $u$ (for *upper*) and $l$ (for *lower*), with energies $E_u$ and $E_l < E_u$. We write $g_u$ and $g_l$ for the statistical weights of these levels, $n_u$ and $n_l$ for their populations, i.e., the number of particles per unit volume in each level. The radiative [10] transitions from one level to the other occur via the **emission** (in the $u \to l$ case) or the **absorption** (in the $l \to u$ case) of a photon whose frequency is

$$\nu_{ul} = \frac{E_u - E_l}{h}. \tag{7.53}$$

There are in fact three processes by which a material particle and a radiation field may be coupled, which Einstein identified as **absorption**, **spontaneous emission**, and **stimulated (or induced) emission**.

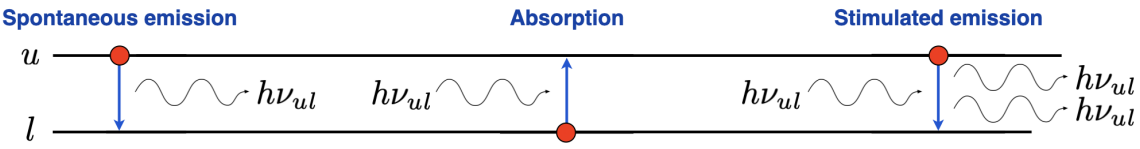



FIGURE 7.3 – The spontaneous emission, absorption, and stimulated emission processes.

10. We ignore collisional transitions here.

#### Spontaneous emission

Spontaneous emission occurs when the atom, initially on the upper level $u$, goes down to level $l$ spontaneously, emitting a photon $h\nu_{ul}$ in the process (Fig. 7.3, left). This happens whether or not a radiation field is present. It is characterized by the **Einstein coefficient of spontaneous emission** $A_{ul}$, which is defined as the probability per unit time that such a transition should occur for a given atom on the upper level. It is therefore expressed in $\mathrm{s}^{-1}$. The number of such transitions per unit volume in a time $\mathrm{d}t$ is then

$$\mathrm{d}n_{ul}^{(e)} = A_{ul} n_u \mathrm{d}t \tag{7.54}$$

#### Absorption

Absorption occurs when the atom, initially on the lower level $l$, goes up to level $u$ by absorbing a photon $h\nu_{ul}$ (Fig. 7.3, center). This requires a radiation field to be present. It is characterized by the **Einstein coefficient of absorption** $B_{lu}$, which is defined so that the probability per unit time that such a transition should occur for a given atom on the lower level is $B_{lu} u_{\nu_{ul}}$. The number of such transitions per unit volume in a time $\mathrm{d}t$ is then

$$\mathrm{d}n_{lu}^{(a)} = B_{lu} u_{\nu_{ul}} n_l \mathrm{d}t \tag{7.55}$$

#### Stimulated emission

Einstein suggested the existence of this third process, by which the atom, initially on the upper level $u$, goes down to level $l$ through the interaction with a photon $h\nu_{ul}$, and emits a second photon in the process, identical to the first one (Fig. 7.3, right). The probability for this process to occur, per unit time, must therefore, like absorption, be proportional to the radiation field at that frequency. The process is characterized by the **Einstein coefficient of stimulated emission** $B_{ul}$, which is defined so that the probability per unit time that such a transition should occur for a given atom on the upper level is $B_{ul} u_{\nu_{ul}}$. The number of such transitions per unit volume in a time $\mathrm{d}t$ is then

$$\mathrm{d}n_{ul}^{(s)} = B_{ul} u_{\nu_{ul}} n_u \mathrm{d}t \tag{7.56}$$

### 7.3.2 Einstein relations

These Einstein coefficients describe processes at the microscopic level. They do not presume of the thermodynamic state of a macroscopic system containing a very large number of such particles in a radiation field. It is therefore possible to suppose that this macroscopic system is in thermal equilibrium and derive relations between the Einstein coefficients that will hold true even outside of such an equilibrium.

At equilibrium, the number of upwards and downwards transitions per unit time must be equal,

$$\mathrm{d}n_{ul}^{(e)} + \mathrm{d}n_{ul}^{(s)} = \mathrm{d}n_{lu}^{(a)} \tag{7.57}$$

This relation may be written in terms of the radiation field's spectral energy density and of the populations of the two levels involved, using the relations defining the Einstein coefficients,

$$\left(A_{ul} + B_{ul} u_{\nu_{ul}}\right) n_u = B_{lu} u_{\nu_{ul}} n_l \tag{7.58}$$

and we reorder this equation to isolate the spectral energy density of the radiation field

$$u_{\nu_{ul}} = \frac{A_{ul} n_u}{B_{lu} n_l - B_{ul} n_u} = \frac{\dfrac{A_{ul}}{B_{ul}}}{\dfrac{B_{lu}}{B_{ul}} \dfrac{n_l}{n_u} - 1} \tag{7.59}$$

Now, at thermal equilibrium, the radiation field obeys Planck's law, and the populations of the two levels are determined by the canonical probabilities of equation (3.9), so

$$\frac{n_l}{n_u} = \frac{g_l}{g_u} \exp\left(-\frac{E_l - E_u}{k_B T}\right) = \frac{g_l}{g_u} \exp\left(\frac{h\nu_{ul}}{k_B T}\right) \tag{7.60}$$

Combining this relation with the expression of the spectral energy density for the radiation we have

$$\frac{8\pi h\nu_{ul}^3}{c^3} \frac{1}{\exp\left(\frac{h\nu_{ul}}{k_B T}\right) - 1} = \frac{\frac{A_{ul}}{B_{ul}}}{\frac{B_{lu}}{B_{ul}} \frac{g_l}{g_u} \exp\left(\frac{h\nu_{ul}}{k_B T}\right) - 1} \tag{7.61}$$

and this relation must hold for any temperature. Consequently, the Einstein coefficients are not independent, and they obey the following two **Einstein relations**

$$g_l B_{lu} = g_u B_{ul} \qquad \frac{A_{ul}}{B_{ul}} = \frac{8\pi h\nu_{ul}^3}{c^3} \tag{7.62}$$

These relations[11] make no reference to the temperature $T$, they are valid even outside of thermal equilibrium and stem from detailed balance relations at the microscopic scale. We note that it suffices to know one of the coefficients to know all three of them. Usually, the coefficient $A_{ul}$ is given in spectroscopic tables[12]. As an order of magnitude, **allowed transitions** have $A_{ul} \sim 10^8\ \mathrm{s}^{-1}$, while **forbidden transitions** have much smaller $A_{ul}$ values (for instance, the HI 21 cm line has $A_{ul} = 2.85\ 10^{-15}\ \mathrm{s}^{-1}$).

### 7.3.3 Kirchhoff's law

Let us consider a material body in thermal equilibrium with a radiation field. The material body absorbs, reflects, and emits photons in order to maintain equilibrium with the field. The isotropy of the black body radiation implies that the amount of radiative energy per unit volume, per frequency interval, and per unit solid angle[13] is

$$\frac{1}{4\pi V}\frac{\mathrm{d}E}{\mathrm{d}\nu} = \frac{u_\omega}{4\pi}\frac{\mathrm{d}\omega}{\mathrm{d}\nu} = \frac{\hbar\omega^3}{2\pi^2 c^3}\frac{1}{e^{\beta\hbar\omega} - 1} = \frac{2h\nu^3}{c^3}\frac{1}{e^{\beta h\nu} - 1} = \frac{B_\nu}{c} \tag{7.63}$$

Now consider a surface $\mathrm{d}A$ on the material body, as in Fig. 7.4, and the amount of radiative energy this surface element receives, in a time $\mathrm{d}t$, carried by photons with frequency in $[\nu, \nu + \mathrm{d}\nu]$ coming from a set of directions within an elementary solid angle $\mathrm{d}\Omega$. This energy is

$$\mathrm{d}e_{\mathrm{inc}} = \frac{B_\nu}{c}\mathrm{d}\nu\mathrm{d}\Omega \times \mathrm{d}A \times c\mathrm{d}t \times \cos\theta \tag{7.64}$$

where $\theta$ is the angle between the incident radiation direction and the normal to the surface element. The product $\mathrm{d}A \times c\mathrm{d}t \times \cos\theta$ represents the volume of the cylinder within which the photons must have been in order for them to reach the surface element between $t$ and $t + \mathrm{d}t$. A fraction $\alpha_\nu(\theta) \leqslant 1$ of this energy, called the **absorptivity**, will be absorbed by the material body, depending on the frequency

11. Note that in some texts slightly different relations are obtained because the Einstein coefficients $B_{ul}$ and $B_{lu}$ are defined with respect to the mean intensity $J_\nu$ of the radiation field rather than with respect to the energy density $u_\nu$.

12. Examples are NIST (`https://www.nist.gov/`), LAMBDA (`https://home.strw.leidenuniv.nl/ moldata/`) or CDMS (`https://cdms.astro.uni-koeln.de/`)

13. We recall that in spherical coordinates, the elementary solid angle is $\mathrm{d}\Omega = \sin\theta\mathrm{d}\theta\mathrm{d}\phi$ and is equal to the ratio of the elementary surface $\mathrm{d}A$ subtended on a sphere to the squared radius $R^2$ of that sphere. The solid angle over all possible directions is $4\pi$.

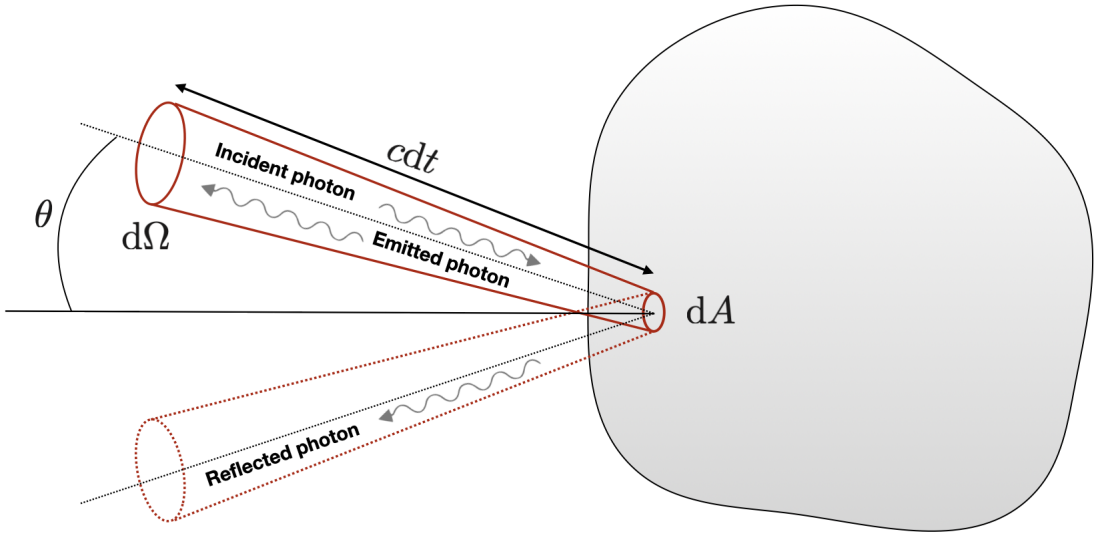



FIGURE 7.4 – Kirchhoff's law

and the incident direction. So the quantity of radiative energy in that frequency interval and elementary solid angle that is effectively absorbed is

$$\mathrm{d}e_{\mathrm{abs}} = \alpha_\nu(\theta) B_\nu \cos\theta \mathrm{d}\nu \mathrm{d}\Omega \mathrm{d}A \mathrm{d}t \tag{7.65}$$

If we assume that whatever reflection there is on the surface does not change either $\theta$ (no scattering, consistent with Snell-Descartes laws) or $\nu$ (no fluorescence), the condition for equilibrium requires that this amount of absorbed radiation be balanced by an emission of photons from the material body, in the same frequency range and the same direction. Let us write the amount of energy emitted in these ranges by a surface $\mathrm{d}A$ of the body during a time interval $\mathrm{d}t$ as

$$\mathrm{d}e_{\mathrm{em}} = j_\nu(\theta) \mathrm{d}\nu \mathrm{d}\Omega \times \cos\theta \mathrm{d}A \times \mathrm{d}t \tag{7.66}$$

where the cosine factor is needed to take into account that the apparent emtting surface, for this direction $\theta$, is $\mathrm{d}A\cos\theta$. This defines $j_\nu(\theta)$. Then equilibrium requires

$$\frac{j_\nu(\theta)}{\alpha_\nu(\theta)} = B_\nu(T) \tag{7.67}$$

so even though the absorption and emission properties of different bodies may vary, we see that, in thermal equilibrium, their ratio is independent of the body, and only depends on the frequency and temperature. This is **Kirchhoff's law**. If scattering or fluorescence are important, it needs to be adapted to include integrations on direction and frequency on each side of the balancing equation $\mathrm{d}e_{\mathrm{abs}} = \mathrm{d}e_{\mathrm{em}}$. A black body is one for which $\alpha_\nu = 1$, absorbing all radiation that reaches it. For non black bodies, a related useful quantity is the **emissivity**[14] $\varepsilon_\nu$, which is defined as the ratio of this body's emission to that of the black body at the same temperature, so Kirchhoff's law may also be stated as

$$\varepsilon_\nu = \frac{j_\nu}{B_\nu} = \alpha_\nu \tag{7.68}$$

14. Note that "emissivity" is a term used for a different yet related concept in radiative transfer.

The emissivity is at most unity, and this is reached only for black-body emission. This is why it is sometimes stated that there cannot exist a physical object emitting more than a black-body. A material body whose emissivity is less than unity but does not depend on frequency is called a **grey body**.

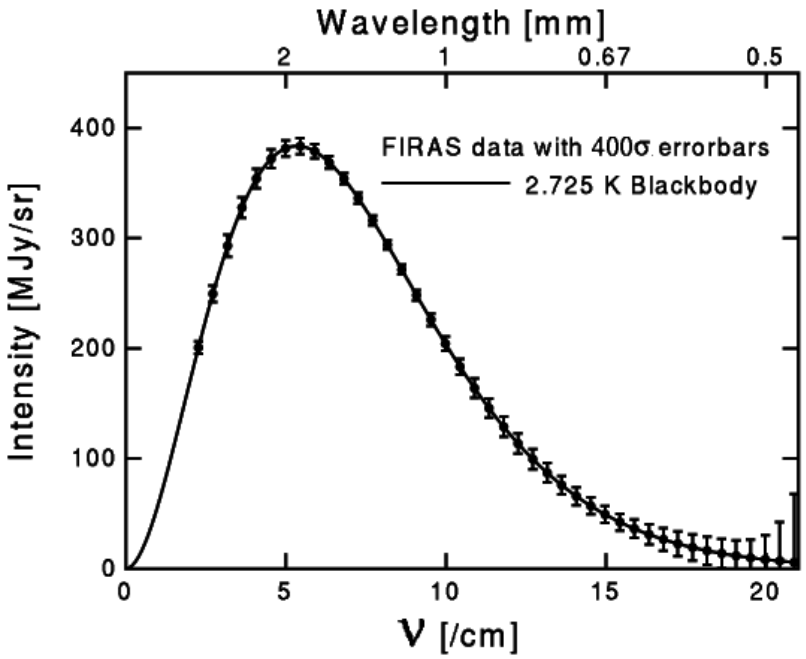



FIGURE 7.5 – Spectrum of the CMB measured by COBE/FIRAS [Mather *et al.*, 1990] .

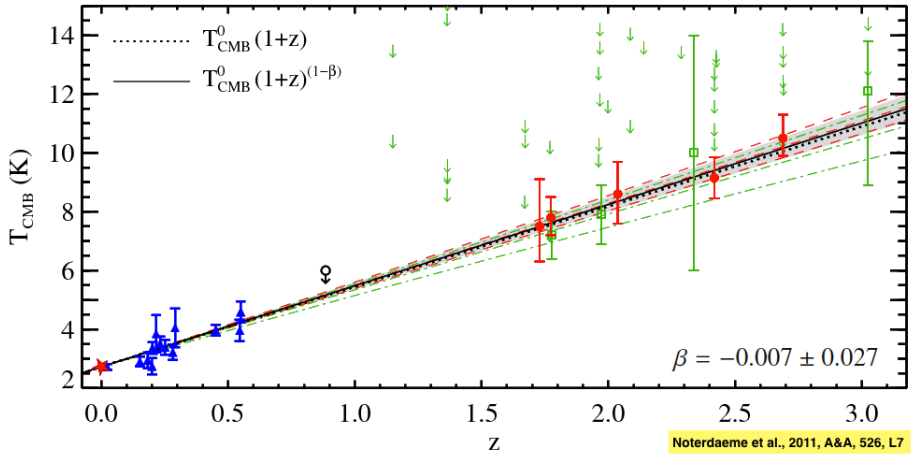



FIGURE 7.6 – CMB temperature as a function of redshift, using molecular line absorption measurements [Noterdaeme *et al.*, A&A, 526, 7, 2011 ].

## 7.4 An example of black-body : the CMB

The **cosmic microwave background (CMB)**, predicted by the works of Alpher, Herman, and Gamow in 1948, and discovered by Penzias and Wilson in 1965 [ApJ, 142, 419], is a near-perfect black body, with no measurable deviation from the Planck law [15], as shown in Fig. 7.5. The error bars on this spectrum measured by COBE/FIRAS in 1990 represent $400\sigma$, which means that the typical uncertainty on each data point in extremely small. The emission peak of $B_\nu$ is around 1.87 mm, corresponding to a frequency of 160 GHz, and a temperature $T_{\rm CMB} = 2.725\,\mathrm{K}$. This is the temperature at the present epoch, and the cosmological expansion suggests that it should vary with the scale factor $a(t)$ of the universe, or equivalently with the redshift $z$, as

$$T_{\rm CMB} \propto \frac{1}{a} \propto 1 + z \tag{7.69}$$

which is consistent with an isentropic evolution of the CMB photon gas, for which $T^3V$ is a constant. This is observed through the measurement of $T_{\rm CMB}(z)$ from molecular absorption lines, such as shown in Fig. 7.6. The origin of this radiation lies in the decoupling of matter and radiation approximately 380 000 years after the Big Bang, when the Universe's temperature fell below $\sim 3000\ \mathrm{K}$. Protons and electrons were then able to recombine, leading to a dramatic increase of the photons' mean free path. These are the ones forming the CMB.

15. There are spectral deviations, that are due to the Sunyaev-Zel'dovich effect, an inverse Compton scattering of the CMB photons on the high-energy electrons present in clusters of galaxies.

# 8

# Equilibrium in an external force field

In this chapter, we consider a macroscopic system of identical particles subject to an external force field. As examples of such a system, we may think of the molecules of a gas in a planetary atmosphere, subject to the gravitational attraction $\vec{g}$ of the planet, or of the electrons in a metal subject to an electrostatic field $\vec{\mathcal{E}}$. Our general hypotheses here are :

— **(H1)** that the force field is constant in time and derives from a potential $\phi$, i.e., $\vec{F} = -\vec{\nabla}\phi$
— **(H2)** that the presence of $\phi(\vec{r})$ does not alter the nature of interparticular interactions
— **(H3)** that $\phi(\vec{r})$ varies appreciably only over macroscopic scales

## 8.1 Equilibrium conditions in an external field

### 8.1.1 Influence of the field on local conditions

#### The system as a collection of sub-systems

Let us consider the system $\mathcal{S}$ as a collection of many mesoscopic-scale disjoint sub-systems $\delta\mathcal{S}_i$, i.e.

$$\mathcal{S} = \bigcup_i \delta\mathcal{S}_i \tag{8.1}$$

The "mesoscopic" condition means that the sub-systems are large enough to contain a very large number of particles, so the methods of statistical physics developed so far may apply, but small enough that over each of these the potential $\phi(\vec{r})$ may be considered approximately uniform. Every such sub-system may exchange energy and particles with the rest of the system, which makes the latter a reservoir of energy and particles for that sub-system. In other words, we can treat each $\delta\mathcal{S}_i$ as if it were in contact with a reservoir that imposes both its temperature $T$ and its chemical potential $\mu$.

#### Grand-canonical description

Writing $E_\ell$ for the energies of the various states of the sub-system in the absence of a field, the second of our general hypotheses **(H2)** ensures that the individual states $|\ell\rangle$ are not modified by the presence of the field, and that this simply alters these energies by adding the corresponding potential energy to each of its particles

$$E_\ell \quad \longrightarrow \quad E'_\ell = E_\ell + N_\ell \phi(\vec{r}) \tag{8.2}$$

From this consideration, it is straightforward to write the probabilities $p'_\ell$ for these microstates

$$p'_\ell = \frac{1}{\Xi'} e^{-\beta[E_\ell + N_\ell \phi(\vec{r}) - \mu N_\ell]} \tag{8.3}$$

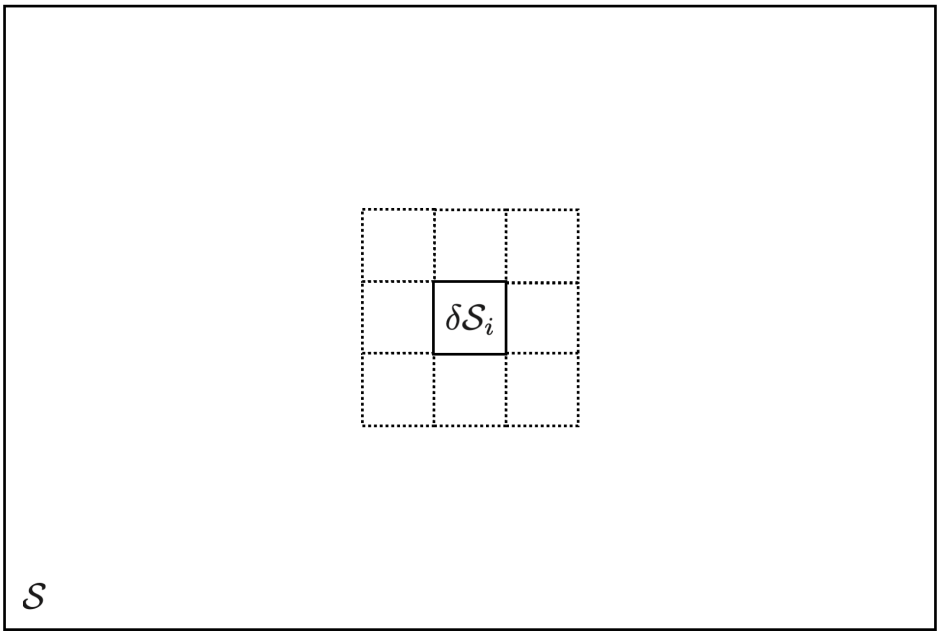



FIGURE 8.1 – A system $\mathcal{S}$ as a collection of sub-systems.

and the corresponding grand-canonical partition function

$$\Xi'(T,\mu,\vec{r}) = \sum_{|\ell\rangle} e^{-\beta[E_\ell + N_\ell\phi(\vec{r}) - \mu N_\ell]} = \sum_{|\ell\rangle} e^{-\beta[E_\ell - \tilde{\mu}(\vec{r})N_\ell]} = \Xi[T,\tilde{\mu}(\vec{r})] \tag{8.4}$$

where the addition of the potential energy translates into a modification of the chemical potential, which becomes a **local chemical potential** dependent on the position in the system $\mathcal{S}$, with

$$\tilde{\mu}(\vec{r}) = \mu - \phi(\vec{r}) \tag{8.5}$$

### 8.1.2 Condition for local equilibrium

Now for the whole system to be considered at equilibrium, it should have a uniform temperature $T$ and chemical potential $\mu$, otherwise parts of the system would spontaneously lose energy or particles, on average, to other parts of $\mathcal{S}$. We therefore must have $T$ uniform, but also

$$\tilde{\mu}(\vec{r}) + \phi(\vec{r}) = \mu = \mathrm{C^{te}} \tag{8.6}$$

The average number of particles in the subsystem located at position $\vec{r}$ is given by

$$\langle N'(\vec{r})\rangle = \sum_{|\ell\rangle} N_\ell p'_\ell = \frac{1}{\Xi'}\sum_{|\ell\rangle} N_\ell e^{-\beta[E_\ell - \tilde{\mu}(\vec{r})N_\ell]} = \frac{1}{\beta}\frac{\partial \ln \Xi'}{\partial \tilde{\mu}} = \frac{1}{\beta}\frac{\partial \ln \Xi}{\partial \tilde{\mu}} \tag{8.7}$$

We recognize here the expression for the average number of particles in the grand-canonical framework if the imposed chemical potential were $\tilde{\mu}$

$$\langle N'(\vec{r})\rangle = \langle N[T,\tilde{\mu}(\vec{r})]\rangle \tag{8.8}$$

For all physical quantities $Y$, except energy, the average value at position $\vec{r}$ is, similarly, equal to that which would be obtained if the imposed chemical potential were $\tilde{\mu}$, i.e.

$$\langle Y'(\vec{r})\rangle = \langle Y\,[T, \tilde{\mu}(\vec{r})]\rangle \tag{8.9}$$

Of course, for the energy, there is the added contribution of the potential energy

$$\langle E'(\vec{r})\rangle = \langle E\,[T, \tilde{\mu}(\vec{r})]\rangle + \langle N\,[T, \tilde{\mu}(\vec{r})]\rangle\phi(\vec{r}) \tag{8.10}$$

### 8.1.3 Interpretations

**Canonical interpretation of the local chemical potential**

Consider a specific sub-system $\delta\mathcal{S}$ of volume $\delta V$ at position $\vec{r}$. Imagine that its number of particles $N$ is fixed, in order to work in the canonical description. Its partition function in the presence of the external field reads

$$Z'(T, \delta V, N) = \sum_{\{|\ell\rangle|N_\ell = N\}} e^{-\beta E'_\ell} = e^{-\beta N\phi(\vec{r})} Z(T, \delta V, N) \tag{8.11}$$

introducing the partition function $Z(T, \delta V, N)$ in the absence of the external field. From this relation, we establish the link between the canonical chemical potential of the subsystem in the absence of the field ($\mu_c$) and in its presence ($\mu'_c$),

$$\mu'_c(T, \delta V, N) = \phi(\vec{r}) + \mu_c(T, \delta V, N) \tag{8.12}$$

This stems from the definition of the Helmholtz free energy and the canonical chemical potential, with

$$\mu'_c = \frac{\partial F'}{\partial N} = -\frac{1}{\beta}\frac{\partial \ln Z'}{\partial N} = -\frac{1}{\beta}\frac{\partial}{\partial N}\left[-\beta N\phi(\vec{r}) + \ln Z\right] = \phi(\vec{r}) + \mu_c \tag{8.13}$$

Now in reality the sub-system is not in the canonical situation, as its number of particles may fluctuate, and its chemical potential is imposed by the rest of the system, but if it is large enough, these fluctuations are small, and the most probable value $\langle N'(\vec{r})\rangle$ in the presence of the field must be such that its canonical chemical potential is equal to that imposed by the rest of the system, i.e.

$$\mu'_c[T, \delta V, \langle N'(\vec{r})\rangle] = \mu \tag{8.14}$$

Now this yields a simple interpretation to $\tilde{\mu}$, since

$$\tilde{\mu}(\vec{r}) = \mu - \phi(\vec{r}) = \mu'_c[T, \delta V, \langle N'(\vec{r})\rangle] - \phi(\vec{r}) = \mu_c[T, \delta V, \langle N'(\vec{r})\rangle] \tag{8.15}$$

so the local chemical potential $\tilde{\mu}$ is just the canonical chemical potential that the sub-system would have *in the absence of the external field*, but for a number of particles equal to the most probable value *in the presence of the field*.

**Physical interpretation of the equilibrium condition**

Take two adjacent sub-systems $\delta\mathcal{S}_1$ and $\delta\mathcal{S}_2$, located at positions $\vec{r}_1$ and $\vec{r}_2$, and suppose that $\phi(\vec{r}_1) > \phi(\vec{r}_2)$. From a mechanical point of view, we expect the force, as the opposite of the gradient of $\phi$, to drag particles from $\delta\mathcal{S}_1$ to $\delta\mathcal{S}_2$. Now the equilibrium condition that $\tilde{\mu}(\vec{r}) + \phi(\vec{r})$ is a constant means that $\tilde{\mu}(\vec{r}_1) < \tilde{\mu}(\vec{r}_2)$. Writing the variation of the Helmholtz free-energy of the joint system $\delta\mathcal{S}_1 \cup \delta\mathcal{S}_2$ due to the sole particle exchange between the two parts, we have

$$\mathrm{d}F = \mathrm{d}F_1 + \mathrm{d}F_2 = [\tilde{\mu}(\vec{r}_1) - \tilde{\mu}(\vec{r}_2)]\,\mathrm{d}N_1 \tag{8.16}$$

and this must be negative in a spontaneous evolution. Therefore, $\mathrm{d}N_1 > 0$ which means that from a statistical point of view, particles tend to leave $\delta\mathcal{S}_2$ for $\delta\mathcal{S}_1$. This is the opposite motion to what the mechanical point of view finds. The equilibrium is due to both of these trends compensating each other.

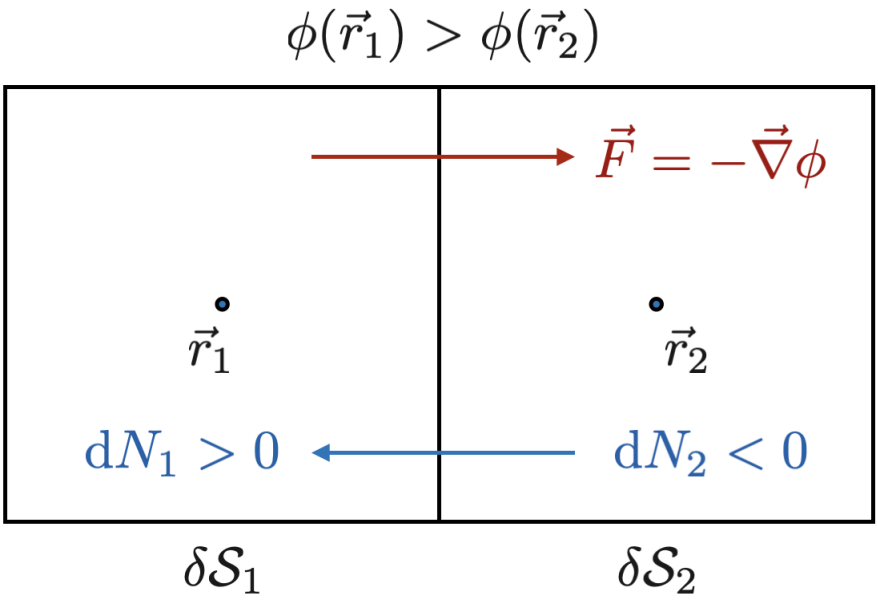



FIGURE 8.2 – Physical interpretation of the equilibrium condition

## 8.2 Examples

### 8.2.1 Simple fluid in a force field

We consider a simple fluid, that is described by three state variables, and for these we choose for convenience, the temperature $T$, pressure $P$, and number of particles $N$. In this system of variables, the only extensive variable is $N$, and therefore the chemical potential, which is intensive, cannot depend on it. In the absence of a field, it only depends on $T$ and $P$, so we write it as $\mu_0(T, P)$. When a non-uniform field $\phi$ is present, the chemical potential varies spatially, but $\tilde{\mu}(\vec{r})$ is, as we saw, the same as $\mu_0$ (in the absence of the field) provided we consider it as a function of the local temperature and pressure. The former remains uniform as we saw, so it must be that the pressure is spatially varying,

$$\tilde{\mu}(\vec{r}) = \mu_0[T, P(\vec{r})] \tag{8.17}$$

The condition for equilibrium then reads

$$\mu_0[T, P(\vec{r})] + \phi(\vec{r}) = \mathrm{C^{te}} \tag{8.18}$$

from which we take the gradient

$$\left(\frac{\partial \mu_0}{\partial P}\right)_T \vec{\nabla} P + \vec{\nabla}\phi = \vec{0} \tag{8.19}$$

The partial derivative in the first term is simply the inverse density [1], and is also a function of position

$$\frac{1}{n(\vec{r})}\vec{\nabla}P + \vec{\nabla}\phi = \vec{0} \tag{8.20}$$

This is the **hydrostatic equilibrium equation**. As an example, we take the case of a classical ideal gas in a uniform gravitational field, so that the problem is one-dimensional, and we have $P = nk_BT$ on the one hand, and $\phi = mgz$ on the other, with $m$ the mass of the particles. The above relation becomes

$$\frac{1}{n}\frac{\mathrm{d}P}{\mathrm{d}z} + mg = \frac{k_BT}{P}\frac{\mathrm{d}P}{\mathrm{d}z} + mg = 0 \tag{8.21}$$

which is easily solved to yield the well-known result that illustrates the Boltzmann factor for gravitational potential energy,

$$P(z) = P_0 \exp\left(-\frac{mgz}{k_BT}\right) \tag{8.22}$$

The **scale height** of this exponential decay is $H = k_BT/mg$, and it may be noted that this decay is quite shallow at room temperatures, with $P$ varying by less that 0.1% for a 10 m ascent. In a more realistic model of the atmosphere, it is necessary to take into account the variation of $T$ with altitude.

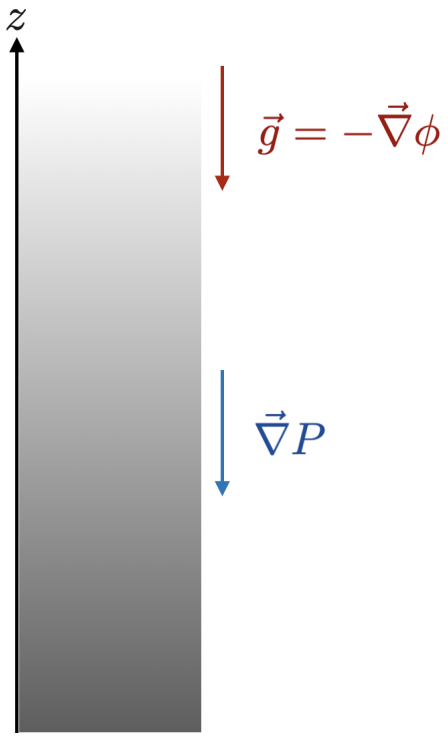



FIGURE 8.3 – Stratification of a simple fluid.

1. Indeed, we have the Gibbs free-energy $G(T, P, N) = N\mu(T, P)$ for a simple fluid, and the required relation stems from the differential expression $\mathrm{d}G = -S\mathrm{d}T + V\mathrm{d}P + \mu\mathrm{d}N$. Using Maxwell's relations, we have

$$\left(\frac{\partial\mu}{\partial P}\right)_T = \left(\frac{\partial V}{\partial N}\right)_P = \frac{1}{n}$$

where the last relation stems from the fact that $V$ and $N$ are extensive variables and so are proportional to each other.

### 8.2.2 Electric equilibrium of a conducting medium

#### Position of the problem

We consider a conducting medium in which positive ions are fixed at certain positions in a lattice, and conduction electrons are free to move around. If no electric field $\vec{\mathcal{E}}_{\text{ext}}$ is applied, the positive and negative charges will balance out over a sufficiently large (mesoscopic) volume, and the net charge density will be null $\rho = 0$. Now, an external electric field $\vec{\mathcal{E}}_{\text{ext}}$ is applied, created by an external charge distribution $\rho_{\text{ext}}(\vec{r})$, and associated to an external electrostatic potential [2] $\varphi_{\text{ext}}$. We recall that

$$\varphi_{\text{ext}}(\vec{r}) = \frac{1}{4\pi\epsilon_0} \iiint \frac{\rho_{\text{ext}}(\vec{r}')}{|\vec{r} - \vec{r}'|} \mathrm{d}^3\vec{r}' \qquad \vec{\mathcal{E}}_{\text{ext}} = -\vec{\nabla}\varphi_{\text{ext}} \tag{8.23}$$

Under this external potential, the electrons are displaced, and potentially the internal charge density $\rho_{\text{ind}}(\vec{r})$ becomes non-zero. This in turns induces an electrostatic potential $\varphi_{\text{ind}}(\vec{r})$ and the associated electric field $\vec{\mathcal{E}}_{\text{ind}}$. The electrons are therefore subject to the total potential $\varphi(\vec{r}) = \varphi_{\text{ext}}(\vec{r}) + \varphi_{\text{ind}}(\vec{r})$.

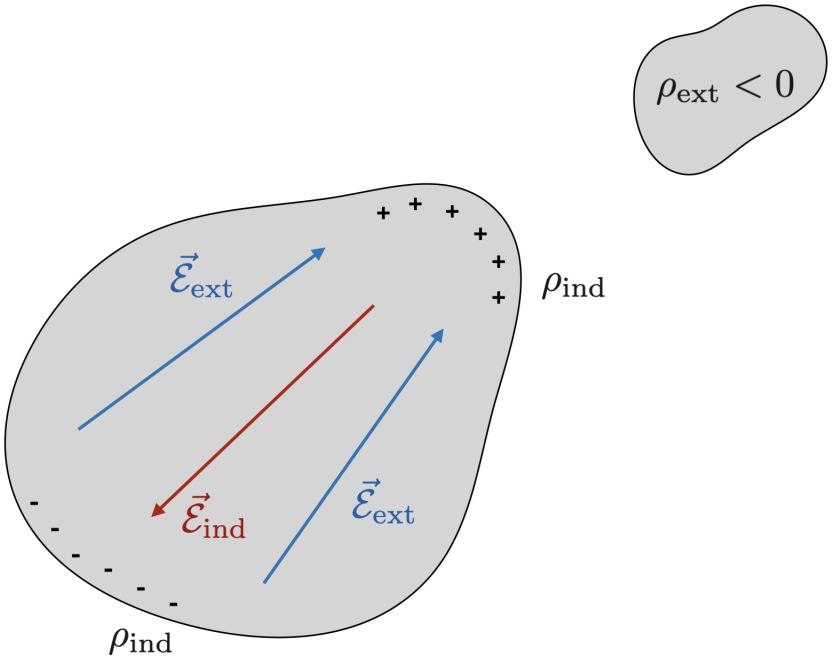



FIGURE 8.4 – Electric equilibrium of a conducting medium.

#### Equations of the problem

We start by writing out the equation for the equilibrium of the electrons in a small volume $\delta V$, introducing the local chemical potential $\tilde{\mu}(\vec{r})$, which must be a function of the canonical variables $T$, $\langle \delta N'(\vec{r}) \rangle$ (the number of electrons in the presence of the external field, that may vary spatially), and $\delta V$. Now, since the chemical potential is an intensive quantity, it must actually be a function of $T$ and $n'(\vec{r}) = \langle \delta N'(\vec{r}) \rangle / \delta V$, so

$$\tilde{\mu}(\vec{r}) + \phi(\vec{r}) = \tilde{\mu}(\vec{r}) - e\varphi(\vec{r}) = \mu_0 \left[T, n'(\vec{r})\right] - e\varphi(\vec{r}) = \mathrm{C}^{\text{te}} \tag{8.24}$$

2. Note that the relationship between $\varphi$ and $\phi$ is that $\phi = q\varphi$ for an electric charge $q$ placed in the electrostatic field.

Now the induced electric charge density is simply due to the excess or lack of electrons [3] compared to the case with no external field

$$\rho_{\mathrm{ind}}(\vec{r}) = -e\left[n'(\vec{r}) - n\right] \tag{8.25}$$

where $n$ is the density in the absence of the external field. The total electrostatic potential must be related to the total charge density through the Poisson equation

$$\Delta\varphi + \frac{1}{\epsilon_0}\rho = \Delta\varphi + \frac{1}{\epsilon_0}\left(\rho_{\mathrm{ext}} + \rho_{\mathrm{ind}}\right) = 0 \tag{8.26}$$

so since $\rho_{\mathrm{ext}}$ is given, the potential is determined by $\rho_{\mathrm{ind}}(\vec{r})$, which itself is determined by $n'(\vec{r})$, and this equation should be used in conjunction with the equilibrium equation (8.24) to determine both $\varphi$ and $n'$ (or $\rho_{\mathrm{ind}}$) consistently. We may assume that the effect of the field on the charge density is small, so that we may write a first order Taylor expansion of the chemical potential as

$$\mu_0\left[T, n'(\vec{r})\right] \approx \mu_0(T,n) + \left(\frac{\partial\mu_0}{\partial n}\right)_T \left[n'(\vec{r}) - n\right] = \mu_0(T,n) - \frac{1}{e}\left(\frac{\partial\mu_0}{\partial n}\right)_T \rho_{\mathrm{ind}}(\vec{r}) \tag{8.27}$$

This shows that in this approximation, $\mu_0\left[T, n'(\vec{r})\right]$ is linear in the induced charge density $\rho_{\mathrm{ind}}$. The equilibrium equation (8.24) then implies that $\varphi$ is also linear in $\rho_{\mathrm{ind}}$, with

$$\varphi(\vec{r}) = \varphi_0 + \frac{1}{e}\left(\frac{\partial\mu_0}{\partial n}\right)_T \left[n'(\vec{r}) - n\right] = \varphi_0 - \frac{1}{e^2}\left(\frac{\partial\mu_0}{\partial n}\right)_T \rho_{\mathrm{ind}}(\vec{r}) \tag{8.28}$$

where $\varphi_0$ is a constant. Finally, taking the Laplacian of that last equation, and combining with the Poisson equation,

$$\Delta\rho_{\mathrm{ind}} - k_0^2\left(\rho_{\mathrm{ext}} + \rho_{\mathrm{ind}}\right) = 0 \qquad\qquad k_0 = \frac{e}{\sqrt{\epsilon_0\left(\frac{\partial\mu_0}{\partial n}\right)_T}} \tag{8.29}$$

The parameter $k_0$ is the inverse of a characteristic length

$$l_0 = \frac{1}{k_0} = \frac{1}{e}\sqrt{\epsilon_0\left(\frac{\partial\mu_0}{\partial n}\right)_T} \tag{8.30}$$

It should be recalled that the partial derivative of $\mu_0$ with respect to $n$ need indeed be positive for the equilibrium to be stable, as found in equation (3.51).

### Characteristic lengths for metals and semiconductors

In the case of a metal, we have

$$\mu_0 = \frac{\hbar^2}{2m}\left(\frac{6\pi^2}{2s+1}\frac{N}{V}\right)^{2/3} = \frac{\hbar^2}{2m}\left(3\pi^2 n\right)^{2/3} \tag{8.31}$$

which is an expression of the kinetic energy at the Fermi level. It is not only valid at $T = 0$ but also at ordinary temperatures due to the very large values [4] of the Fermi temperature $T_F$. We then have for $l_0$ an expression that is called the **Thomas-Fermi length**, which is of order 1 Å.

$$l_0 = l_{\mathrm{TF}} = \frac{\hbar\pi^{2/3}}{e}\sqrt{\frac{\epsilon_0}{m}}(3n)^{-1/6} \tag{8.32}$$

3. Remember that ions are assumed to be fixed, their density does not change in the presence of the field.
4. The density of charges in a metal is of the order $10^{22}\,\mathrm{cm}^{-3}$, leading to Fermi temperatures in excess of $10^4\,\mathrm{K}$.

In the case of a semi-conductor, the density of free electrons is much smaller than in a metal, so that this may be treated as a classical (rather than fermionic) ideal gas. The chemical potential in that case is given by equation (3.90), i.e., [5]

$$\mu_0 = -k_B T \ln\left(\frac{z}{N}\right) = -k_B T \ln\left[\frac{2V}{N\Lambda^3}\right] = -k_B T \ln\left[2\frac{V}{N}\frac{(2\pi m k_B T)^{3/2}}{h^3}\right] \tag{8.33}$$

which may be written more concisely to underline the role of the density $n$ as

$$\mu_0 = -k_B T \ln\left[\frac{\alpha(T)}{n}\right] \qquad \alpha(T) = 2\frac{(2\pi m k_B T)^{3/2}}{h^3} \tag{8.34}$$

The characteristic length is then the **Debye length**, which appears also in the context of space plasmas, which are also low-density media,

$$l_0 = l_{\mathrm{D}} = \sqrt{\frac{\epsilon_0 k_B T}{n e^2}} \tag{8.35}$$

### Screening

Consider the case where the external field is created by a single, outside charge $q_0$ located at the origin of the coordinate system. In that case, $\rho_{\mathrm{ext}}(\vec{r}) = q_0\delta(\vec{r})$ and if we consider that charge to be sufficiently far away from the metal's surface, we can assume that its effects are spherically symmetric. Consequently, equation (8.29) reads

$$\frac{1}{r}\frac{\mathrm{d}^2}{\mathrm{d}r^2}\left(r\rho_{\mathrm{ind}}\right) - k_0^2\rho_{\mathrm{ind}} = k_0^2 q_0 \delta(\vec{r}) \tag{8.36}$$

and the right-hand side vanishes outside of the origin. The solution to this is a linear combination of exponential terms, only one of which is physical, so

$$\rho_{\mathrm{ind}}(\vec{r}) = \frac{A}{r}e^{-k_0 r} \qquad \varphi(\vec{r}) = \frac{A'}{r}e^{-k_0 r} \tag{8.37}$$

The constants may be obtained by reinjecting into the Poisson equation, finding the asymptotic behaviour as $r \to 0$, and comparing that result with the Gauss theorem for a single charge since when we consider a point very close to the charge $q_0$ the potential should be dominated by that part. Eventually, we obtain

$$\varphi(\vec{r}) = \frac{q_0}{4\pi\epsilon_0 r}e^{-k_0 r} \tag{8.38}$$

which, compared to the external potential created by the single outside charge, is attenuated very rapidly (on a scale $\sim l_0$) because this charge attracts opposite ones towards it, so that for a charge further out, the net apparent charge is diminished. This is a phenomenon called **screening**.

5. We recall that $s = 1/2$ for electrons, so the spin state factor is $2s + 1 = 2$.

# 9 Magnetism

The **magnetic properties** of materials provide a rich topic to apply the methods of statistical physics. In this chapter, we will discuss the microscopic models for these magnetic properties, distinguishing between systems in which elements are independent (leading to **paramagnetism** and **diamagnetism**), and those where interactions are essential, leading to **ferromagnetism**.

## 9.1 Generalities

In the presence of a magnetic field $\vec{B}$, most materials acquire a **macroscopic magnetic moment** $\vec{M}$, that is in general parallel to $\vec{B}$. The average magnetic moment per unit volume is called the **magnetization**

$$\vec{\mathcal{M}} = \frac{\langle \vec{M} \rangle}{V} \tag{9.1}$$

Depending on whether the magnetization is parallel or anti-parallel to the magnetic field, the material is said to be **paramagnetic** or **diamagnetic**. Paramagnetism only occurs when the elements of the system have permanent microscopic magnetic moments $\vec{\mu}$, that align in the applied field $\vec{B}$. An example of a paramagnetic compound is $\mathrm{O}_2$. Diamagnetism is caused by the change of electron motion in the atoms or molecules of the material when a magnetic field is applied. These **induced magnetic moments** are such that they oppose the applied field [1] and so the magnetization is anti-parallel to the field. Examples of diamagnetic compounds are bismuth, silver, and carbon. In general, paramagnetism is larger than diamagnetism [2]. In ferromagnetic materials, microscopic magnetic moments are locally ordered even in the absence of an applied field. These properties are sketched in Fig. 9.1
The response of the material to the applied field is quantified by the **magnetic susceptibility**, defined by

$$\chi = \lim_{B \to 0} \frac{\partial \mathcal{M}}{\partial B} \tag{9.2}$$

where $\vec{\mathcal{M}} = \mathcal{M}\vec{e}_z$ if $\vec{B} = B\vec{e}_z$. Since magnetization has dimensions of an electric current per unit length ($\mathrm{A} \cdot \mathrm{m}^{-1}$), it has the same dimensions as a magnetic field divided by the magnetic permeability

1. This is **Lenz's law**.
2. In a diluted gas of free electrons, the two effects exist, and are of the same order of magnitude.

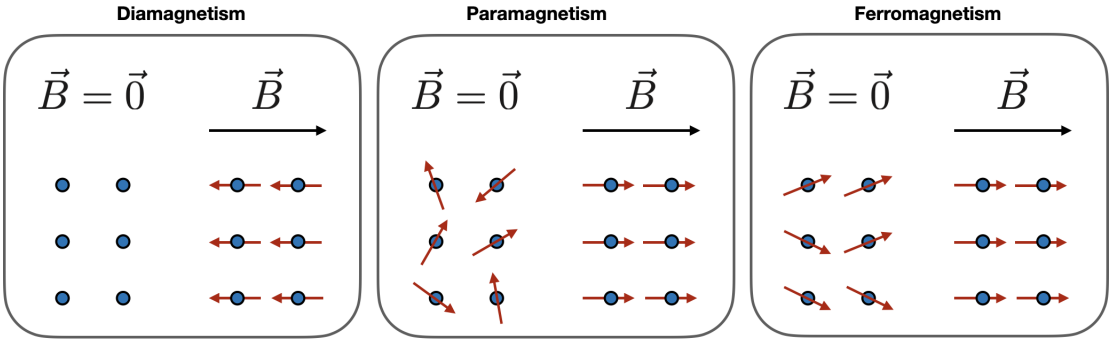



FIGURE 9.1 – Schematic behaviours of diamagnetic, paramagnetic, and ferromagnetic materials with and without an applied magnetic field.

$\mu_0$, as can be shown from the Maxwell-Ampère law [3]. Thus, the susceptibility has the same dimension as the inverse of $\mu_0$, and is usually given in terms of the dimensionless $\hat{\chi} = \mu_0 \chi$. Obviously, $\chi > 0$ for paramagnetic materials and $\chi < 0$ for diamagnetic ones [4]. The treatment of these systems (and of ferromagnetic ones later) may be done in the canonical framework, based on the fact that the energy associated to a magnetic moment $\vec{\mu}$ in a magnetic field $\vec{B}$ is

$$\varepsilon = -\vec{\mu} \cdot \vec{B} \tag{9.3}$$

Thus, if a system of volume $V$ consisting of $N$ magnetic moments $\{\vec{\mu}_i\}$, in contact with a thermostat at temperature $T$, acquires a total magnetic moment $\vec{M}_\ell$ when in state $|\ell\rangle$, the associated energy is $E_\ell = -\vec{M}_\ell \cdot \vec{B}$ and one may easily find [5] that the magnetization and susceptibility are

$$\mathcal{M} = \frac{k_B T}{V} \frac{\partial \ln Z}{\partial B} = -\frac{1}{V} \frac{\partial F}{\partial B} \qquad \chi = -\frac{1}{V} \lim_{B \to 0} \frac{\partial^2 F}{\partial B^2} \tag{9.5}$$

These expressions hold whether we are considering a diamagnetic, paramagnetic, or ferromagnetic material. We will now explore the microscopic models of these in turn, focusing first on paramagnetic systems, then discussing at length the topic of ferromagnetism. We will not discuss the case of diamagnetism.

## 9.2 Paramagnetism

### 9.2.1 Hamiltonian

Let us consider a paramagnetic material of volume $V$, at temperature $T$, to which we apply a uniform magnetic field $\vec{B} = B\vec{e}_z$. The material consists of $N$ permanent magnetic moments $\vec{\mu}_i$, that may be related to the respective angular momenta [6] $\vec{J}_i$ through

$$\vec{\mu}_i = g\mu_B \vec{J}_i \tag{9.6}$$

3. A magnetic moment is expressed in $\mathrm{A} \cdot \mathrm{m}^2$, so a magnetization is indeed in $\mathrm{A} \cdot \mathrm{m}^{-1}$. The Maxwell-Ampère law gives that $[B]/L = \mu_0[j] = \mu_0 \cdot \mathrm{A} \cdot \mathrm{m}^{-2}$ and thus $[B]/\mu_0 = \mathrm{A} \cdot \mathrm{m}^{-1}$.

4. We have typically $|\hat{\chi}| \sim 10^{-5} - 10^{-6}$ for diamagnetic liquids ands solids, and $\hat{\chi} \sim 10^{-2}$ for paramagnetic ones.

5. It suffices to write the average magnetic moment $\langle \vec{M} \rangle$ using the canonical probabilities $p_\ell$ of Eq.(3.9), i.e.,

$$\mathcal{M} = \frac{1}{V} \sum_{|\ell\rangle} p_\ell M_\ell = \frac{1}{ZV} \sum_{|\ell\rangle} M_\ell e^{-\beta E_\ell} = \frac{1}{ZV} \sum_{|\ell\rangle} M_\ell e^{\beta M_\ell B} = \frac{1}{ZV} \frac{\partial Z}{\partial(\beta B)} = \frac{k_B T}{V} \frac{\partial \ln Z}{\partial B} \tag{9.4}$$

6. These are the total momenta, comprising the orbital and spin contributions. They are expressed in units of $\hbar$, and are therefore dimensionless.

where $g$ is the **Landé factor**, of order unity, and $\mu_B$ is the **Bohr magneton**, with

$$\mu_B = \frac{e\hbar}{2m_e} \approx 9.274\,10^{-24}\,\mathrm{A}\cdot\mathrm{m}^2 \tag{9.7}$$

We already tackled such a system in previous chapters, when considering $N$ atoms, each with a spin $s = 1/2$ and associated magnetic moment $\vec{\mu}$ that can point either in the direction of $\vec{B}$ ($|+\rangle$ state) or in the opposite direction ($|-\rangle$ state), i.e., $\mu_z = \pm\mu$.

The **Hamiltonian** of the system, neglecting interactions between magnetic moments, is then

$$H = \sum_{i=1}^{N} \varepsilon_i = -\sum_{i=1}^{N} \vec{\mu}_i \cdot \vec{B} = -g\mu_B \vec{B} \cdot \sum_{i=1}^{N} \vec{J}_i = -g\mu_B B \sum_{i=1}^{N} J_i \tag{9.8}$$

where $J_i \in \{-J, -J+1, \ldots, J-1, J\}$ are the eigenvalues of the projection of each $\vec{J}_i$ on the $Oz$ axis.

### 9.2.2 The case of the two-level system $J = 1/2$

For simplicity, we will consider the important case of the two-level system $J = 1/2$ from now on. The formulas for the general case are given in section 9.2.3.

#### Partition function and Helmholtz free-energy

Since the particles are independent and distinguishable, as they are at specific discrete sites, the canonical partition function is $Z = z^N$, where $z$ is the partition function of a single particle. It is straightforwardly computed as

$$Z = z^N = \left[\sum_{J=-1/2}^{1/2} \exp\left(\frac{g\mu_B B J}{k_B T}\right)\right]^N = \left[2\cosh\left(\frac{g\mu_B B}{2k_B T}\right)\right]^N \tag{9.9}$$

From this, the Helmholtz free-energy is

$$F = -Nk_B T \ln\left[2\cosh\left(\frac{g\mu_B B}{2k_B T}\right)\right] \tag{9.10}$$

#### Magnetization and susceptibility

We obtain the magnetization from the general formula (9.5), which gives [7]

$$\mathcal{M} = \frac{Ng\mu_B}{2V} \tanh\left(\frac{g\mu_B B}{2k_B T}\right) \tag{9.12}$$

The **saturation magnetization** is then, not surprinsigly,

$$\mathcal{M}_\infty = \frac{Ng\mu_B}{2V} \tag{9.13}$$

7. We have

$$\mathcal{M} = -\frac{1}{V}\frac{\partial F}{\partial B} = \frac{1}{V}\frac{\partial}{\partial B}\left(Nk_B T \ln\left[2\cosh\left(\frac{g\mu_B B}{2k_B T}\right)\right]\right) = \frac{Nk_B T}{V}\frac{\frac{g\mu_B}{2k_B T}\sinh\left(\frac{g\mu_B B}{2k_B T}\right)}{\cosh\left(\frac{g\mu_B B}{2k_B T}\right)} \tag{9.11}$$

Note that we already found this relationship, in Eq. (2.49), by writing explicitly the average magnetic moment using the probabilities of the microstates, i.e.,

$$\mathcal{M} = \frac{1}{V}\left(n_+\mu - n_-\mu\right) = \frac{\mu N}{V}\left(\frac{e^{\beta\mu B} - e^{-\beta\mu B}}{e^{\beta\mu B} + e^{-\beta\mu B}}\right) = \frac{\mu N}{V}\tanh\left(\frac{\mu B}{k_B T}\right) \tag{9.14}$$

and noting that we have $\mu = g\mu_B/2$ in this case. The magnetization has the behaviour sketched in Fig. 9.2, which is simple to understand : at very low temperatures or very high fields, the spins align themselves on the field, while at low field or high temperatures, the thermal fluctuations make the orientations of the various spins completely random, resulting in no net magnetization.

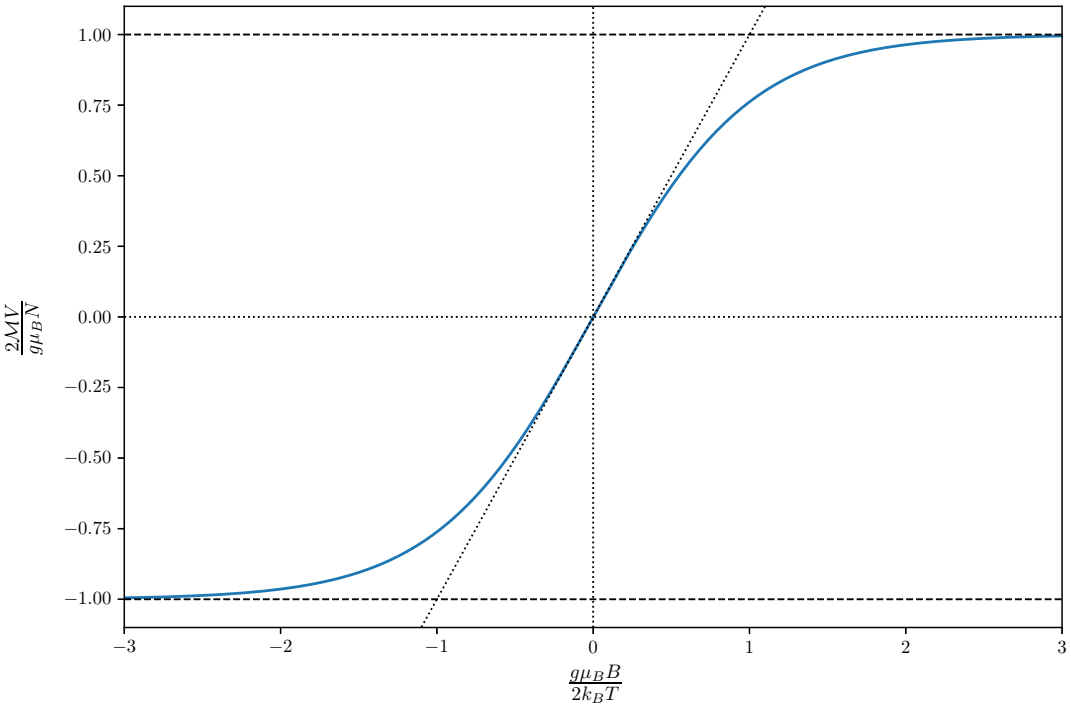



FIGURE 9.2 – Magnetization of the perfect paramagnetic crystal.

The magnetic susceptibility is then [8]

$$\chi = \lim_{B\to 0}\frac{\partial\mathcal{M}}{\partial B} = \lim_{B\to 0}\left[\frac{Ng\mu_B}{2V}\times\frac{g\mu_B}{2k_B T}\frac{1}{\cosh^2\left(\frac{g\mu_B B}{2k_B T}\right)}\right] = \frac{N(g\mu_B)^2}{4Vk_B T} \tag{9.15}$$

We find **Curie's law** that the magnetic susceptibility of a paramagnetic material is inversely proportional to the temperature, with $\chi = \mathcal{C}/T$ and $\mathcal{C}$ being the **Curie constant**.

8. We recall that $(\tanh x)' = 1/\cosh^2 x$.

**Energy and heat capacity**

The average energy is given by Eq. (3.27), which gives [9]

$$\langle E\rangle = -\frac{Ng\mu_B B}{2}\tanh\left(\frac{g\mu_B B}{2k_B T}\right) \tag{9.17}$$

which, unsurprisingly, is equal to $-V\mathcal{M}B$. The associated heat capacity is then [10]

$$C = Nk_B\left(\frac{g\mu_B B}{2k_B T}\right)^2 \frac{1}{\cosh^2\left(\frac{g\mu_B B}{2k_B T}\right)} \tag{9.19}$$

It should be stressed that this is the energy and heat capacity associated to the magnetic energy only, and that one should add to this any other type of energy relevant to the system under study, in particular that associated to the vibrational degrees of freedom in the solid (see Chapter 10).

**Entropy**

The entropy $S$ is computed from the general fomula $F = \langle E\rangle - TS$, yielding

$$S = Nk_B\left\{\ln\left[2\cosh\left(\frac{g\mu_B B}{2k_B T}\right)\right] - \frac{g\mu_B B}{2k_B T}\tanh\left(\frac{g\mu_B B}{2k_B T}\right)\right\} \tag{9.20}$$

As it should, and as can be seen in Fig. 9.3, entropy goes to zero as $T \to 0$ or $B \to \infty$, since in that case all magnetic moments are aligned and the system is therefore perfectly determined. In the inverse limit $B/T \to 0$, we get $S = Nk_B \ln 2 = k_B \ln(2^N)$ which is simply expressing that the $2^N$ states of the system are equally likely. An important application of the fact that entropy is actually a function of the ratio $B/T$ is **adiabatic demagnetization**, a process by which a slow, reversible demagnetization of an isolated paramagnetic system leads to significant cooling, reaching temperatures [11] of order $\sim 1\,\mathrm{mK}$

### 9.2.3 Formulas for the general case of $J$

In the general case of a quantized angular momentum $J$ (either integer or half-integer), we have the canonical partition function [12]

$$Z = \left[\sum_{J_i=-J}^{J}\exp\left(\frac{g\mu_B B J_i}{k_B T}\right)\right]^N = \left\{\frac{\sinh\left[\frac{g\mu_B B}{k_B T}\left(J+\frac{1}{2}\right)\right]}{\sinh\left(\frac{g\mu_B B}{2k_B T}\right)}\right\}^N \tag{9.21}$$

9. We have

$$\langle E\rangle = -\frac{\partial \ln Z}{\partial\beta} = -N\frac{\partial}{\partial\beta}\left\{\ln 2 + \ln\left[\cosh\left(\frac{g\mu_B B\beta}{2}\right)\right]\right\} = -N\frac{\frac{g\mu_B B}{2}\sinh\left(\frac{g\mu_B B\beta}{2}\right)}{\cosh\left(\frac{g\mu_B B\beta}{2}\right)} \tag{9.16}$$

10. We have

$$C = \frac{\partial\langle E\rangle}{\partial T} = -\frac{Ng\mu_B B}{2}\frac{1}{\cosh^2\left(\frac{g\mu_B B}{2k_B T}\right)}\times\left(-\frac{g\mu_B B}{2k_B T^2}\right) \tag{9.18}$$

11. One has to start from an already very low temperature of order $\sim 1\,\mathrm{K}$.
12. The reader is invited to check that the case $J = 1/2$ is properly recovered from these more general formulas.

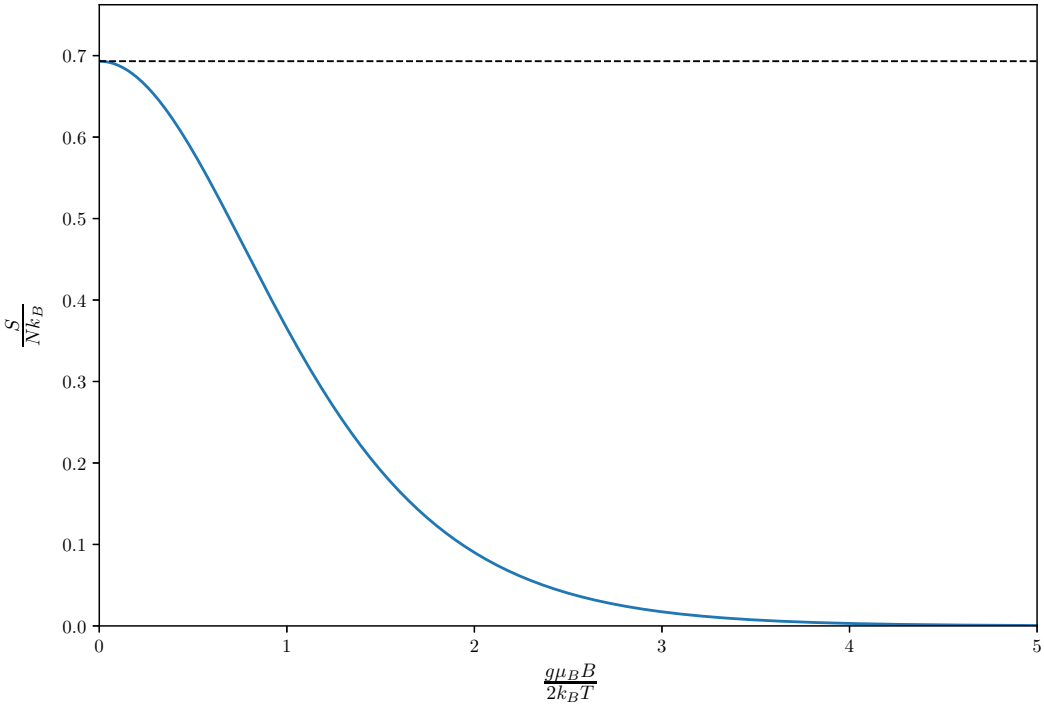



FIGURE 9.3 – Entropy of the perfect paramagnetic crystal, normalized to $Nk_B$, as a function of the $B/T$ ratio, also properly normalized. The horizontal dashed line corresponds to $S/(Nk_B) = \ln 2$, reached in the limit $T \to 0$ or $B \to \infty$.

from which we get the Helmholtz free-energy

$$F = -Nk_B T \ln \left\{ \frac{\sinh\left[\frac{g\mu_B B}{k_B T}\left(J + \frac{1}{2}\right)\right]}{\sinh\left(\frac{g\mu_B B}{2k_B T}\right)} \right\} \tag{9.22}$$

and from that the magnetization, shown in Fig. 9.4 for various values of $J$,

$$\mathcal{M} = \frac{Ng\mu_B J}{V} B_J\left(\frac{g\mu_B JB}{k_B T}\right) = \mathcal{M}_\infty B_J\left(\frac{g\mu_B JB}{k_B T}\right) \tag{9.23}$$

where $B_J$ is the **Brillouin function of order** $J$ defined by

$$B_J(x) = \frac{2J+1}{2J}\coth\left(\frac{2J+1}{2J}x\right) - \frac{1}{2J}\coth\left(\frac{x}{2J}\right) \tag{9.24}$$

The magnetic susceptibility also obeys Curie's law

$$\chi = \frac{N(g\mu_B)^2 J(J+1)}{3Vk_B T} \tag{9.25}$$

The mean energy is given by the following formula, which may also be found from $\langle E\rangle = -V\mathcal{M}B$,

$$\langle E\rangle = -Ng\mu_B BJB_J\left(\frac{g\mu_B JB}{k_B T}\right) \tag{9.26}$$

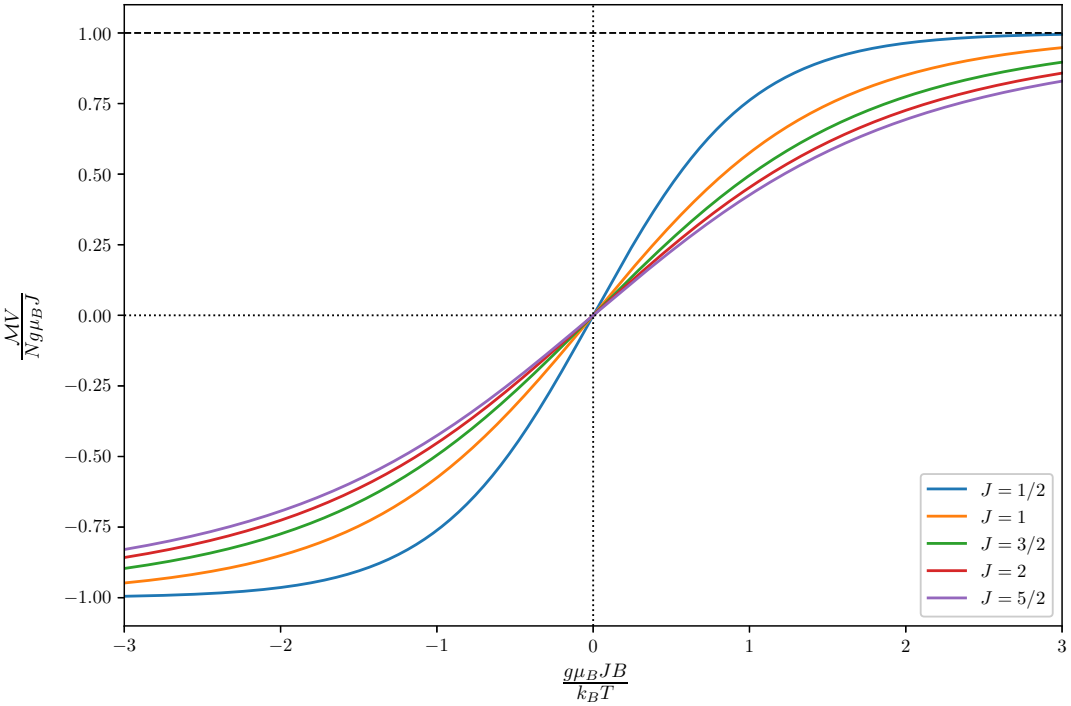



FIGURE 9.4 – Magnetization of the perfect paramagnetic crystal for various cases of $J$.

The heat capacity and entropy may be computed as an exercise. Note that for the latter it is recovered that $S \to 0$ in the limit $B/T \to \infty$, and that in the other limit we get $S \to Nk_B \ln(2J+1)$ as it should since there are $2J+1$ equiprobable states for each of the $N$ particles.

## 9.3 The phenomenon of ferromagnetism

The phenomenon of **ferromagnetism** is the property of some materials[13] to act as permanent magnets, i.e., to possess a **spontaneous** magnetization $\vec{\mathcal{M}}$, even in the absence of any applied field $\vec{B}$.

### 9.3.1 Experimental results

#### Magnetization

It is found experimentally that heating a ferromagnetic material - in the absence of an external magnetic field - above a critical temperature, known as the **Curie temperature**, makes it lose its ferromagnetic properties, rendering the material **paramagnetic**. This **phase transition** is of the second order in the classification due to Ehrenfest, so that there is no latent heat and the two phases cannot coexist. For instance, the Curie temperature for iron is $T_\mathrm{C} \approx 1043\,\mathrm{K} \approx 770\,^\circ\mathrm{C}$. When measuring the magnetization $\mathcal{M}$ as a function of the applied (uniform) magnetic field $\vec{B} = B\vec{e}_z$, a ferromagnetic material exhibits the behaviour shown on the left of Fig. 9.5, while a paramagnetic material's behaviour is shown on the right. The spontaneous magnetization appears as the non-zero value of $\mathcal{M}$ when $B = 0$.

13. Iron, nickel, and cobalt are examples, but compounds such as $\mathrm{MnSb}$ are also ferromagnetic.

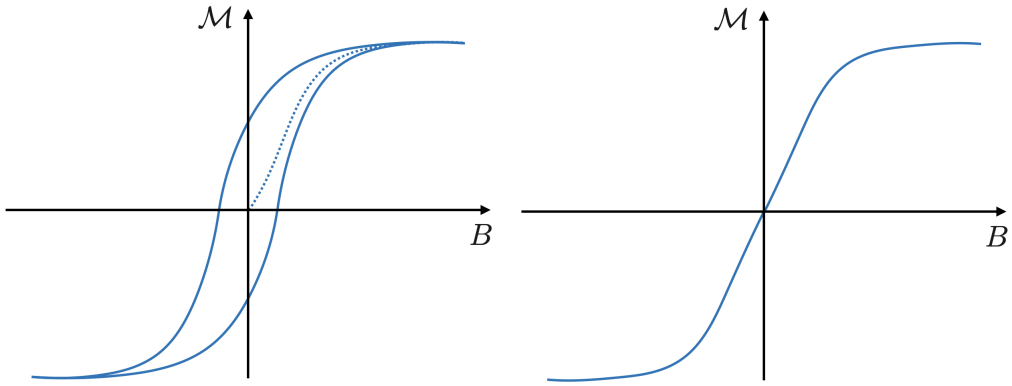



FIGURE 9.5 – Magnetization vs. applied field for a ferromagnetic (left) and paramagnetic (right) material. On the left side, the curve of first magnetization is shown as a dashed line.

Magnetization only vanishes when a sufficiently large opposing field (the **coercive field**) is applied. It appears also that magnetization is not a sole function of the applied field but of the history of the system (a phenomenon called **hysteresis**).

The suggestion that magnetization is non-zero even in the absence of a magnetic field would suggest that these materials should produce a magnetic field. Yet, iron and other materials are in fact found to be unmagnetized as a whole, because a bulk of these materials is in fact composed of many small domains, called **Weiss magnetic domains**, within which the magnetization is indeed non-zero, but different from that of the surrounding domains, as shown in Fig. 9.6. The boundaries of the domains are called **Bloch walls**. On the whole, the vector sum of the magnetizations of the different domains is essentially zero. When applying an increasing magnetic field, the magnetizations of the various domains start to align. When considering a single domain, the magnetization vs. applied field curve has the behaviour sketched in Fig. 9.7, where the case $T < T_\mathrm{C}$ displays a jump from $-\mathcal{M}_0(T)$ to $\mathcal{M}_0(T)$ for $B = 0$, while the curve is continuous for $T > T_\mathrm{C}$. This zero-field magnetization $\mathcal{M}_0(T)$ is shown as a function of temperature in Fig. 9.11, showing that it goes to a finite limit when $T \to 0$, decreases for an increasing temperature, and vanishes - as it should - at the Curie temperature.

### Properties in the vicinity of the critical point

In the vicinity of the critical point, i.e., for $|T - T_\mathrm{C}| \ll T_\mathrm{C}$, the thermodynamic properties of the system exhibit power-law scaling behaviours. In particular, the zero-field magnetization is

$$\mathcal{M}_0(T) \propto \begin{cases} 0 & \text{for } T > T_\mathrm{C} \\ (T_\mathrm{C} - T)^\beta & \text{for } T < T_\mathrm{C} \end{cases} \tag{9.27}$$

where $\beta$ is a fractional exponent to be determined. Another example of such a behaviour is the heat capacity at zero field

$$C(T) \propto \begin{cases} (T - T_\mathrm{C})^{-\alpha} & \text{for } T > T_\mathrm{C} \\ (T_\mathrm{C} - T)^{-\alpha'} & \text{for } T < T_\mathrm{C} \end{cases} \tag{9.28}$$

where both $\alpha$ and $\alpha'$ are positive exponents. A third example is the magnetic susceptibility

$$\chi(T) = \lim_{B \to 0} \frac{\partial \mathcal{M}}{\partial B} \propto \begin{cases} (T - T_\mathrm{C})^{-\gamma} & \text{for } T > T_\mathrm{C} \\ (T_\mathrm{C} - T)^{-\gamma'} & \text{for } T < T_\mathrm{C} \end{cases} \tag{9.29}$$

It turns out experimentally that whatever the (ferromagnetic) material, $\beta \approx 0.34 - 0.38$, $\alpha \approx \alpha' \approx 0.1$, and $\gamma \approx \gamma' \approx 1.3 - 1.4$. Other second-order phase transitions display critical behaviours with similar exponents, leading to an idea of **universal exponents**, the origin of which needs to be explained.

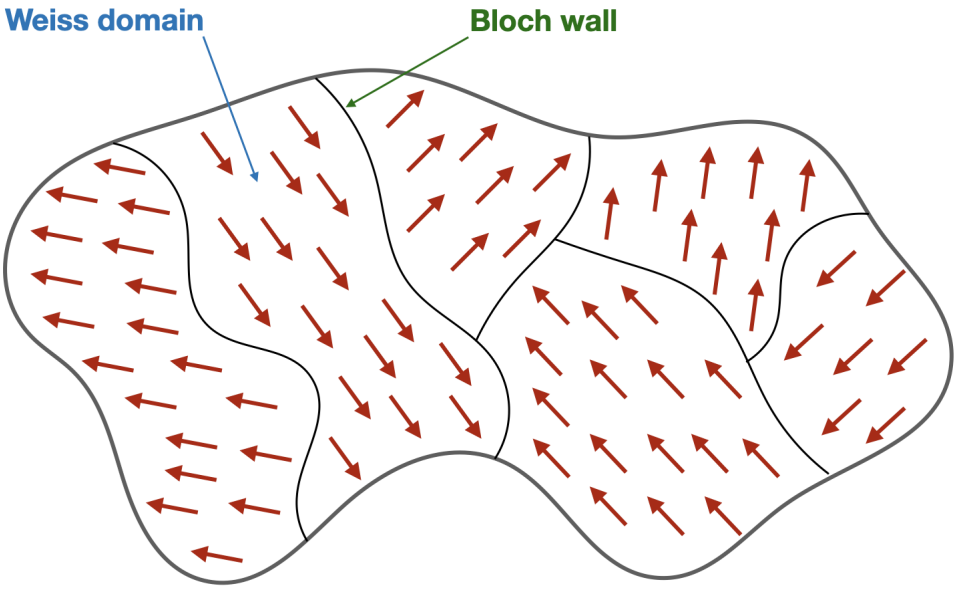



FIGURE 9.6 – Weiss magnetic domains, separated by Bloch walls.

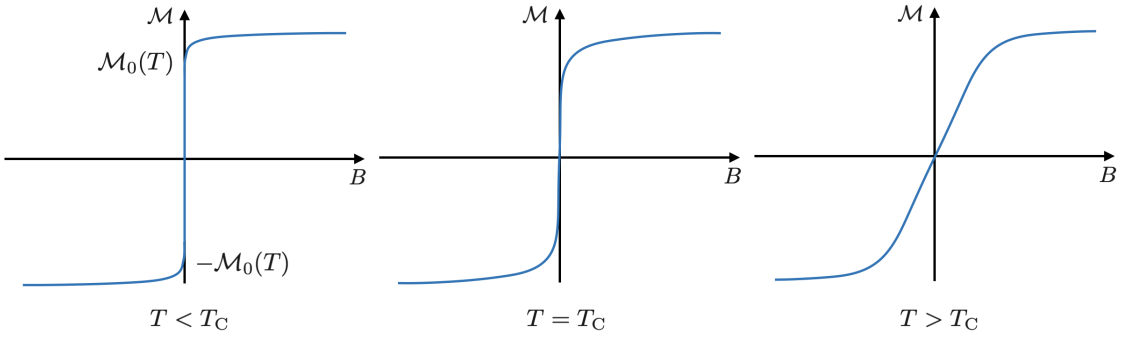



FIGURE 9.7 – Magnetization vs. applied field for a single Weiss domain.

### 9.3.2 The problem to solve

**The Heisenberg Hamiltonian**

Let us consider a single Weiss domain of a ferromagnetic material of volume $V$, at temperature $T$, to which we apply a uniform magnetic field $\vec{B}_0 = B_0\vec{e}_z$. The magnetic moment of the material being $M$, the magnetization is $\mathcal{M} = M/V$. The question we ask is to find the expression of $\mathcal{M}(B_0, T)$ as a function of the applied magnetic field and temperature. For a ferromagnetic system, the essential part left out of the "paramagnetic" Hamiltonian (9.8) is the interaction between spins, as it is clear that each magnetic moment's orientation must be influenced by the local magnetic field created by the neighbouring magnetic moments [14]. This interaction is modelled by the **Heisenberg Hamiltonian**

14. In fact, the origin of the interaction is more complex, as is explained later, and the effect mentioned here is in fact included in the applied field $\vec{B}_0$.

$$H_{\mathrm{H}} = -g\mu_B \vec{B}_0 \cdot \sum_{i=1}^{N} \vec{S}_i - K \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j \tag{9.30}$$

where $K > 0$ is a coupling coefficient [15] and the vectors $\vec{S}_i$ are the spins [16] that have a length $1/2$ and project onto $\vec{e}_z$ either as $+1/2$ or $-1/2$. In the second term, the sum is over the pairs of neighbouring sites, noted $\langle i, j \rangle$.

**The origin of the coupling**

It is important to realize that the origin of this interaction is quantum in nature [17]. Indeed, if one were to compute the energy due to classical dipole-dipole interaction, one would obtain

$$E_{\mathrm{int}} = \frac{\mu_0}{4\pi r^3} \left[ \vec{\mu}_1 \cdot \vec{\mu}_2 - \frac{3}{r^2} \left( \vec{\mu}_1 \cdot \vec{r} \right) \left( \vec{\mu}_2 \cdot \vec{r} \right) \right] \tag{9.31}$$

for two magnetic moments $\vec{\mu}_1$ and $\vec{\mu}_2$ separated by $\vec{r}$. In order of magnitude, this gives

$$E_{\mathrm{int}} \sim \frac{\mu_0 \mu_B^2}{4\pi r^3} \sim \frac{10^{-7} \times 10^{-46}}{10^{-30}} \sim 10^{-23}\,\mathrm{J} \sim 10^{-4}\,\mathrm{eV} \tag{9.32}$$

This order of magnitude corresponds to a temperature of about $1\,\mathrm{K}$, so it is much too weak to explain a phase transition occurring at $\sim 1000\,\mathrm{K}$. The interaction at stake is in reality an **exchange interaction** (Fig. 9.8), due to the combination of the Pauli exclusion principle and the Coulomb interaction (that is of order $\sim 1\,\mathrm{eV}$). In short, the wavefunction $\psi = \phi \otimes \chi$ of, say, a pair of electrons must be fully antisymmetric, so that if its spin part $\chi$ is symmetric (parallel spins), its spatial part $\phi$ must be antisymmetric, resulting in a null probability that the electrons should be at the same location, and a very small probability that they could be close in space. The Coulomb interaction energy, which is positive, is then quite small, much smaller than what it is when the spin wavefunction is antisymmetric (opposite spins), because the spatial wavefunction must then be symmetric and electrons can be spatially much closer, resulting in a strong electrostatic repulsion, i.e. a strong positive interaction energy. Thus, antiparallel spins result in a much larger interaction energy than parallel spins, with an interaction energy of the order $1\,\mathrm{eV}$. This explains the sign of $K$ and the magnitude of the transition temperature.

## 9.4 The mean field approximation for ferromagnetism

### 9.4.1 Principle of the approximation

The **mean field approximation** consists in viewing each magnetic moment as interacting with an average field resulting from the combination of the applied external field and the microscopic fields created by all other magnetic moments. The solution will then have to be found through a **self-consistency** condition. To that effect, we consider the part $H_i$ of the Heisenberg Hamiltonian associated to the $i^{\mathrm{th}}$ spin,

$$H_i = -g\mu_B \vec{B}_0 \cdot \vec{S}_i - K \vec{S}_i \cdot \sum_{j \in v(i)} \vec{S}_j = -g\mu_B \vec{S}_i \cdot \left[ \vec{B}_0 + \frac{K}{g\mu_B} \sum_{j \in v(i)} \vec{S}_j \right] = -g\mu_B \vec{S}_i \cdot \left( \vec{B}_0 + \vec{B}_m \right) \tag{9.33}$$

15. A more general expression would involve coupling coefficients $K_{ij}$ for all pairs of spins, that would decrease as the distance between the two sites increases.

16. We use the notation $\vec{S}$ here instead of $\vec{J}$ used earlier, but the meaning is the same.

17. In fact, quantum physics is essential even to explain paramagnetism and diamagnetism, as a theorem by J.H. van Leeuwen (1911) shows that a purely classical treatment of magnetism leads to a perfect compensation of the paramagnetic and diamagnetic effects, resulting in a net zero magnetization.

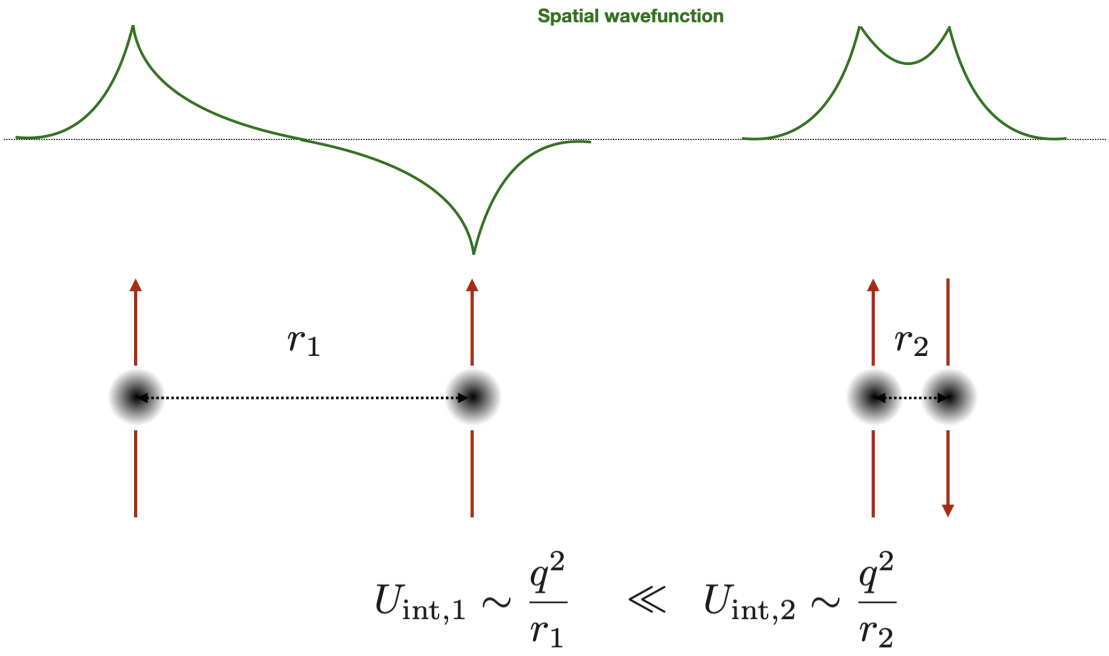



FIGURE 9.8 – Exchange interaction

where $v(i)$ is the ensemble of neighbouring points of $i$. In this expression, $\vec{B}_m$ is the **molecular field** due to all [18] other magnetic moments. On each site, the effective magnetic field is $\vec{B}_i = \vec{B}_0 + \vec{B}_m$, and this is *a priori* a different field at each site since it depends on the magnetic moments at the neighbouring points. The mean field approximation consists in assuming that the effective field is the same for all sites, equal to the mean field

$$\vec{B}_i \approx \vec{B}_{\text{eff}} = \langle \vec{B}_i \rangle = \vec{B}_0 + \frac{K}{g\mu_B} \sum_{j \in v(i)} \langle \vec{S}_j \rangle \tag{9.34}$$

The magnetization, which is an average already, is then given by

$$\vec{\mathcal{M}} = \frac{g\mu_B}{V} \sum_i \langle \vec{S}_i \rangle = \frac{Ng\mu_B}{V} \langle \vec{S}_i \rangle \tag{9.35}$$

The mean field that results is, for a system where each site has $p$ nearest neighbours,

$$\vec{B}_{\text{eff}} = \vec{B}_0 + \frac{K}{g\mu_B} \sum_{j \in v(i)} \langle \vec{S}_j \rangle = \vec{B}_0 + \frac{pVK}{N(g\mu_B)^2} \vec{\mathcal{M}} = \vec{B}_0 + \lambda \vec{\mathcal{M}} \tag{9.36}$$

We assume that both $\vec{B}_0$ and $\vec{\mathcal{M}}$ are along $\vec{e}_z$, with $\vec{\mathcal{M}} = \mathcal{M}\vec{e}_z$. Using the results for the paramagnetic crystal, we find the magnetization as

$$\mathcal{M} = \frac{Ng\mu_B}{2V} \tanh\left[ \frac{g\mu_B}{2k_B T} \left( B_0 + \frac{pVK}{N(g\mu_B)^2} \mathcal{M} \right) \right] \tag{9.37}$$

This equation is to be solved self-consistently, and we have no validity condition for this approximation.

18. This is limited to the nearest neighbours, in this case.

### 9.4.2 Helmholtz free-energy in the mean field approximation

#### Mean-field Hamiltonian

In the mean field approximation, the Hamiltonian is then given by that of a system of $N$ independent spins, to within an additive constant $A$, i.e.,

$$H_{\text{mean}} = -g\mu_B \vec{B}_{\text{eff}} \cdot \sum_i \vec{S}_i + A \tag{9.38}$$

The reason why this constant needs to be introduced is that by replacing the actual field $\vec{B}_i = \vec{B}_0 + \vec{B}_m$ by the effective field $\vec{B}_{\text{eff}}$ we make an approximation and therefore we cannot simply replace one by the other without allowing for some discrepancy between the two expressions. To determine the constant $A$, we write that the mean value of this Hamiltonian should be the same as the mean value of the Heisenberg Hamiltonian, i.e.,

$$-g\mu_B \vec{B}_{\text{eff}} \cdot \sum_i \langle \vec{S}_i \rangle + A = -g\mu_B \vec{B}_0 \cdot \sum_{i=1}^{N} \langle \vec{S}_i \rangle - K \sum_{\langle i,j \rangle} \langle \vec{S}_i \cdot \vec{S}_j \rangle \tag{9.39}$$

Since the mean field approximation essentially neglects correlations between spins, we have

$$\langle \vec{S}_i \cdot \vec{S}_j \rangle = \langle \vec{S}_i \rangle \cdot \langle \vec{S}_j \rangle = \langle \vec{S} \rangle^2 \tag{9.40}$$

and therefore, since there are $Np/2$ pairs of nearest neighbours [19],

$$-Ng\mu_B \vec{B}_{\text{eff}} \cdot \langle \vec{S} \rangle + A = -Ng\mu_B \vec{B}_0 \cdot \langle \vec{S} \rangle - K \frac{Np}{2} \langle \vec{S} \rangle^2 \tag{9.41}$$

Replacing $\langle \vec{S} \rangle$ by its expression as a function of magnetization, derived from $\vec{\mu} = g\mu_B \vec{S}$,

$$\langle \vec{S} \rangle = \frac{V}{Ng\mu_B} \vec{\mathcal{M}} \tag{9.42}$$

and inserting that into the expression of the effective field

$$\vec{B}_{\text{eff}} = \vec{B}_0 + \frac{pVK}{N(g\mu_B)^2} \vec{\mathcal{M}} = \vec{B}_0 + \frac{pK}{g\mu_B} \langle \vec{S} \rangle \tag{9.43}$$

we have that

$$-Ng\mu_B \left[ \vec{B}_0 + \frac{pK}{g\mu_B} \langle \vec{S} \rangle \right] \cdot \langle \vec{S} \rangle + A = -Ng\mu_B \vec{B}_0 \cdot \langle \vec{S} \rangle - K \frac{Np}{2} \langle \vec{S} \rangle^2 \tag{9.44}$$

leading straightforwardly to

$$A = \frac{KNp}{2} \langle \vec{S} \rangle^2 = \frac{KNp}{2} \left( \frac{V\mathcal{M}}{Ng\mu_B} \right)^2 \tag{9.45}$$

and to a mean-field Hamiltonian

$$H_{\text{mean}} = -g\mu_B \vec{B}_{\text{eff}} \cdot \sum_i \vec{S}_i + \frac{KNp}{2} \left( \frac{V\mathcal{M}}{Ng\mu_B} \right)^2 \tag{9.46}$$

This constant $A$ depends on the magnetization that we seek, so it is not just an arbitrary constant that could be done away with. In the mean field approximation, properties of the system are computed as if $\mathcal{M}$ was fixed externally, but it is in fact an internal variable that adjusts itself in order to minimize the Helmholtz free-energy $F$. In that situation, we should recover the self-consistency condition (9.37).

19. There are $N$ sites, each of which has $p$ nearest beighbours. The factor $1/2$ corrects for the fact that this computation leads to counting each pair of sites twice.

**Helmholtz free-energy**

In order to determine $F$, we write the mean-field Hamiltonian as a sum of independent terms

$$H_{\text{mean}} = \sum_{i=1}^{N} H_i \qquad H_i = -g\mu_B \vec{S}_i \cdot \left(\vec{B}_0 + \lambda\vec{\mathcal{M}}\right) + \frac{Kp}{2}\left(\frac{V\mathcal{M}}{Ng\mu_B}\right)^2 \tag{9.47}$$

which means that the canonical partition function is simply $Z_{\text{mean}} = \prod z_i = z^N$ with

$$z = e^{-\beta\epsilon_+} + e^{-\beta\epsilon_-} \qquad \text{with} \qquad \epsilon_\pm = \pm\frac{1}{2}g\mu_B\left(B_0 + \lambda\mathcal{M}\right) + \frac{Kp}{2}\left(\frac{V\mathcal{M}}{Ng\mu_B}\right)^2 \tag{9.48}$$

combining all this into the expression of the canonical partition function, we have

$$Z_{\text{mean}} = \left\{2\exp\left[-\frac{pK}{2k_BT}\left(\frac{V\mathcal{M}}{Ng\mu_B}\right)^2\right]\cosh\left[\frac{g\mu_B}{2k_BT}\left(B_0 + \lambda\mathcal{M}\right)\right]\right\}^N \tag{9.49}$$

The Helmholtz free-energy is then

$$F_{\text{mean}} = -k_BT\ln Z_{\text{mean}} = \frac{KNp}{2}\left(\frac{V\mathcal{M}}{Ng\mu_B}\right)^2 - Nk_BT\ln\left\{2\cosh\left[\frac{g\mu_B}{2k_BT}\left(B_0 + \lambda\mathcal{M}\right)\right]\right\} \tag{9.50}$$

Should we seek the magnetization, we would find $\mathcal{M}$ such that $F_{\text{mean}}$ is minimum, and that would give the self-consistency equation (9.37) [20], but what is more interesting is to study $F_{\text{mean}}$ to decide upon the stability of equilibrium solutions, as we shall see.

### 9.4.3 Properties in zero field

**Spontaneous magnetization**

When the applied field is null, ($B_0 = 0$), the magnetization should obey the self-consistency equation

$$\frac{\mathcal{M}}{\mathcal{M}_\infty} = \tanh\left(\frac{pK}{4k_BT}\frac{\mathcal{M}}{\mathcal{M}_\infty}\right) \tag{9.54}$$

where $\mathcal{M}_\infty$ is the magnetization at zero temperature (**saturation magnetization**), reached when all the spins in the crystal are aligned,

$$\mathcal{M}_\infty = \frac{Ng\mu_B}{2V} \tag{9.55}$$

20. We have indeed, searching for this minimum of the Helmholtz free-energy,

$$\frac{\partial F_{\text{mean}}}{\partial\mathcal{M}} = KNp\left(\frac{V}{Ng\mu_B}\right)^2\mathcal{M} - Nk_BT\frac{\sinh\left[\frac{g\mu_B}{2k_BT}\left(B_0 + \lambda\mathcal{M}\right)\right]}{\cosh\left[\frac{g\mu_B}{2k_BT}\left(B_0 + \lambda\mathcal{M}\right)\right]}\frac{g\mu_B\lambda}{2k_BT} = 0 \tag{9.51}$$

implying that

$$KNp\left(\frac{V}{Ng\mu_B}\right)^2\mathcal{M} = \frac{Ng\mu_B}{2}\frac{pVK}{N(g\mu_B)^2}\tanh\left[\frac{g\mu_B}{2k_BT}\left(B_0 + \lambda\mathcal{M}\right)\right] \tag{9.52}$$

and therefore, as expected

$$\mathcal{M} = \frac{Ng\mu_B}{2V}\tanh\left[\frac{g\mu_B}{2k_BT}\left(B_0 + \lambda\mathcal{M}\right)\right] \tag{9.53}$$

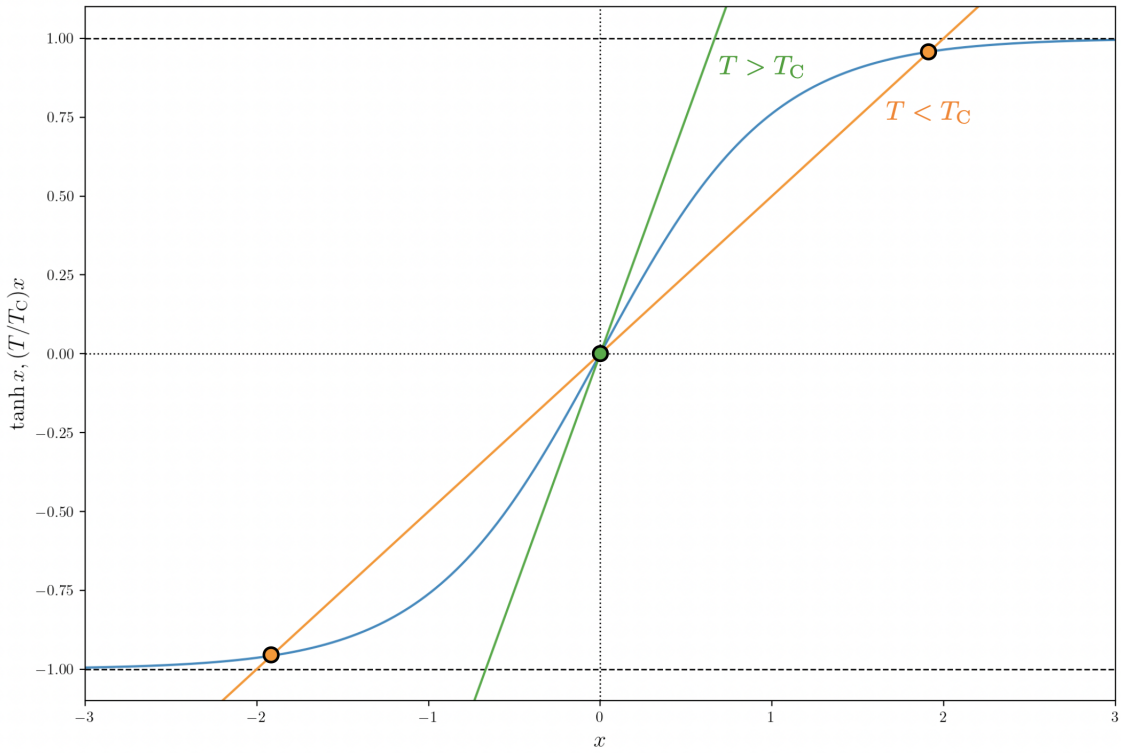



FIGURE 9.9 – Graphical determination of the magnetization in zero applied field. Two cases are shown, for $T < T_C$ in orange and $T > T_C$ in green.

This equation may be solved graphically as shown in Fig. 9.9, introducing the Curie temperature

$$T_C = \frac{pK}{4k_B} \tag{9.56}$$

and plotting [21] both $x \mapsto \tanh x$ and $x \mapsto (T/T_C)x$. It appears that $\mathcal{M} = 0$ is a solution in all cases, but that if $T < T_C$, there are two other possible, non-zero solutions for the magnetization, i.e., $\pm\mathcal{M}_0(T)$. This transition corresponds to a change in the shape of the Helmholtz free-energy [22], as shown in Fig. 9.10, which is computed from the general equation (9.50) in the case $B_0 = 0$,

$$F_{\text{mean}} = Nk_B \left\{ \frac{T_C}{2} \left( \frac{\mathcal{M}}{\mathcal{M}_\infty} \right)^2 - T \ln \left[ 2 \cosh \left( \frac{T_C}{T} \frac{\mathcal{M}}{\mathcal{M}_\infty} \right) \right] \right\} \tag{9.58}$$

Above the Curie temperature, $\mathcal{M} = 0$ is the only, stable solution, while below that threshold it becomes unstable and the non-zero solutions $\pm\mathcal{M}_0(T)$ are stable. Of course, this only gives the magnitude of the magnetization, its direction is chosen randomly in each Weiss domain. The spontaneous magnetization as a function of temperature is shown in Fig. 9.11.

21. In these expressions, we set $x = (T_C/T)(\mathcal{M}/\mathcal{M}_\infty)$.
22. The expression below is found from the previous one by noticing that

$$\frac{KNp}{2} = 2NT_C \qquad \frac{V}{Ng\mu_B} = \frac{1}{2\mathcal{M}_\infty} \tag{9.57}$$

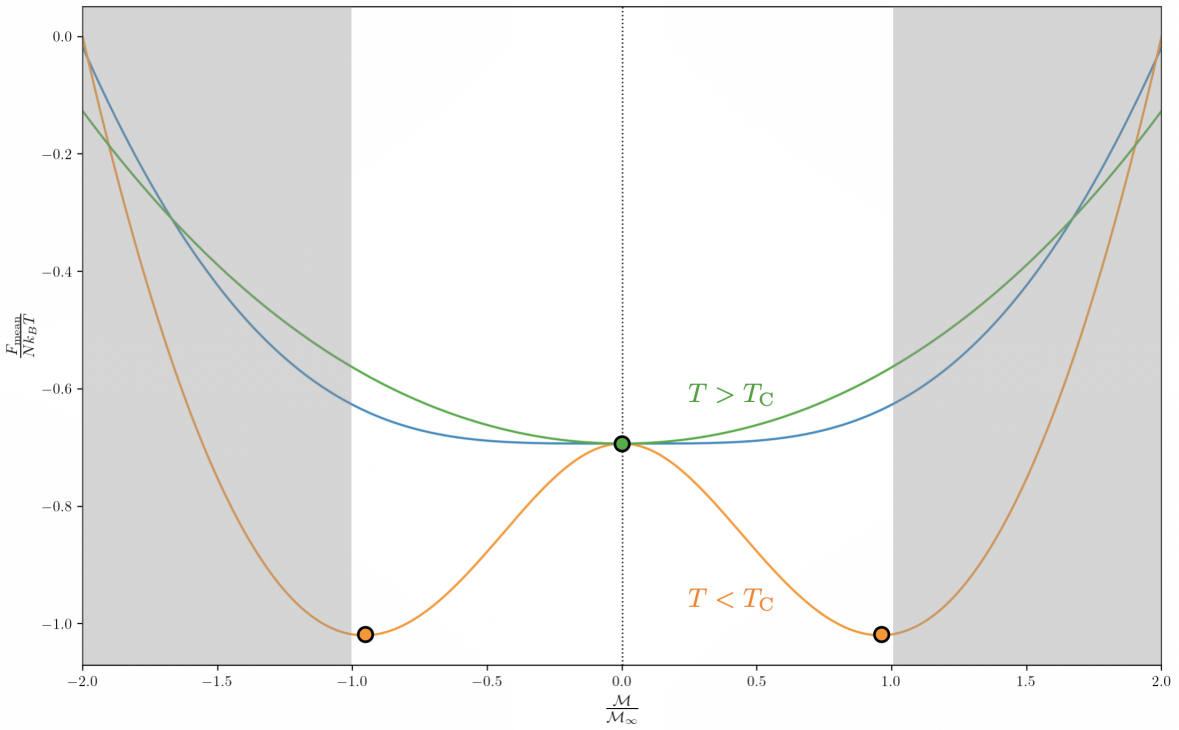



FIGURE 9.10 – Helmholtz free-energy as a function of magnetization, in zero applied field and mean-field approximation, in the cases $T > T_C$ (green), $T = T_C$ (blue), and $T < T_C$ (orange). The grey-shaded areas are beyond the saturation magnetization and are therefore unphysical.

### Properties in the vicinity of the critical temperature

In the vicinity of $T_C$, the spontaneous magnetization tends to zero, and it is straightforward to determine that, in this limit [23]

$$\frac{\mathcal{M}_0}{\mathcal{M}_\infty} \approx \sqrt{3\left(1 - \frac{T}{T_C}\right)} \tag{9.63}$$

23. We use the Taylor expansion of the hyperbolic tangent $\tanh x \approx x - (x^3/3)$ for $x \ll 1$, so that

$$\frac{\mathcal{M}}{\mathcal{M}_\infty} = \frac{T_C}{T}\frac{\mathcal{M}}{\mathcal{M}_\infty} - \frac{1}{3}\left(\frac{T_C}{T}\frac{\mathcal{M}}{\mathcal{M}_\infty}\right)^3 \tag{9.59}$$

and therefore

$$1 - \frac{T_C}{T} = -\frac{1}{3}\left(\frac{T_C}{T}\right)^3\left(\frac{\mathcal{M}}{\mathcal{M}_\infty}\right)^2 \tag{9.60}$$

so that

$$\left(\frac{\mathcal{M}}{\mathcal{M}_\infty}\right)^2 = 3\left(\frac{T}{T_C}\right)^2\left(1 - \frac{T}{T_C}\right) \approx 3\left(1 - \frac{T}{T_C}\right) \tag{9.61}$$

It is also easy to determine that when $T \to 0$, we have, from $\tanh x = (1 - e^{-2x})/(1 + e^{-2x}) \approx 1 - 2e^{-2x}$ for $x \gg 1$, that

$$\frac{\mathcal{M}_0}{\mathcal{M}_\infty} \approx 1 - 2\exp\left(-\frac{2T_C}{T}\right) \tag{9.62}$$

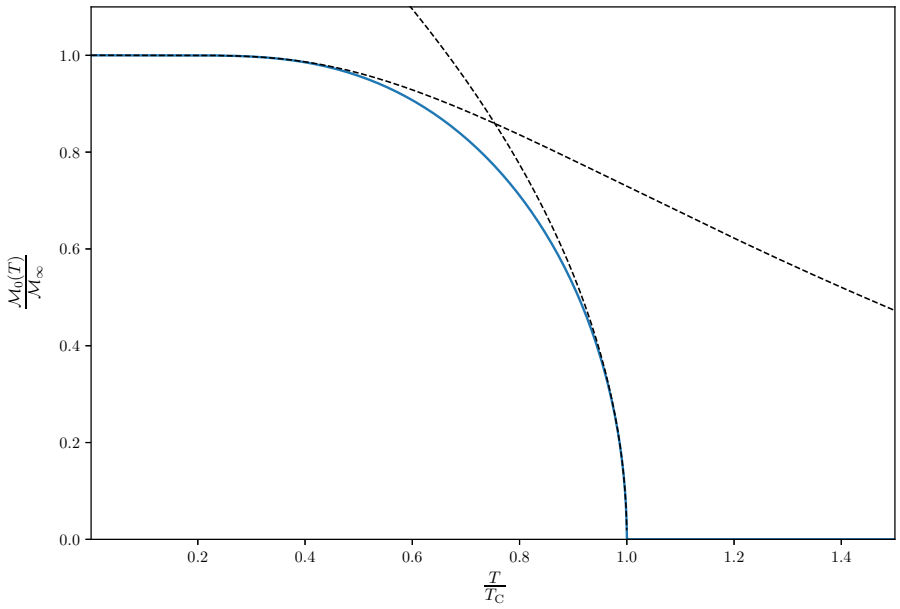



FIGURE 9.11 – Spontaneous magnetization $\mathcal{M}_0(T)$ as a function of temperature. The dashed lines represent the limits discussed in the text, for $T \to 0$ and $T \to T_C$.

The mean energy of the system is given by

$$\langle E \rangle = -\frac{\partial \ln Z_{\text{mean}}}{\partial \beta} = -\frac{NKp}{8}\left(\frac{\mathcal{M}_0}{\mathcal{M}_\infty}\right)^2 \tag{9.64}$$

so it is zero above the Curie temperature, and when approaching it from below, it behaves as

$$\langle E \rangle \approx -\frac{3NKp}{8}\left(1 - \frac{T}{T_C}\right) \tag{9.65}$$

From this we get the heat capacity $C$, that is also null above $T_C$ and positive below, increasing as $T$ increases

$$C = \frac{\partial \langle E \rangle}{\partial T} = -\frac{NKp}{4\mathcal{M}_\infty^2}\mathcal{M}_0 \frac{\mathrm{d}\mathcal{M}_0}{\mathrm{d}T} \tag{9.66}$$

leading to a **discontinuity** at $T = T_C$, with

$$C\left(T_C^-\right) - C\left(T_C^+\right) = \frac{3}{8}\frac{NKp}{T_C} = \frac{3}{2}Nk_B \tag{9.67}$$

This discontinuity makes the ferromagnetic transition a **second-order phase transition**.

### 9.4.4 Properties in non-zero field

In this case, a graphical solution is also possible, as shown in Fig. 9.12, involving the functions [24]

$$x \mapsto \tanh x \qquad\qquad x \mapsto \frac{T}{T_\mathrm{C}} x - \frac{g\mu_B B_0}{2k_B T_\mathrm{C}} \tag{9.69}$$

It may be shown, as an exercise, that if $T > T_\mathrm{C}$ there is always a single intersection point, corresponding to the sole minimum of the Helmholtz free-energy. The magnetization increases as $B_0$ increases, reaching $\mathcal{M}_\infty$ for $B_0 \to \infty$. The magnetic susceptibility is found for $T > T_\mathrm{C}$ in the limit $B_0 \to 0$ as

$$\chi = \lim_{B_0 \to 0} \frac{\partial \mathcal{M}}{\partial B_0} = \frac{N}{V} \left(\frac{g\mu_B}{2}\right)^2 \frac{1}{k_B (T - T_\mathrm{C})} \tag{9.70}$$

This is the same behaviour as a perfect paramagnetic crystal, replacing $T$ by $T - T_\mathrm{C}$, forming the **Curie-Weiss law**. Below the Curie temperature, there may be either one or three intersection points, depending on the strength of the applied field :

— If $B_0 > B_0^0(T)$, there is only one intersection point. In that case, the Helmholtz free-energy only has one minimum.

— If $B_0 < B_0^0(T)$, there are three intersection points. In that case, the Helmholtz free-energy has two minima of different depths, separated by a local maximum.

The magnetic susceptibility just below the Curie temperature yields

$$\chi = \lim_{B_0 \to 0} \frac{\partial \mathcal{M}}{\partial B_0} = \frac{N}{2V} \left(\frac{g\mu_B}{2}\right)^2 \frac{1}{k_B (T_\mathrm{C} - T)} \tag{9.71}$$

that is the same expression as just above the Curie temperature, except for a factor $-1/2$.

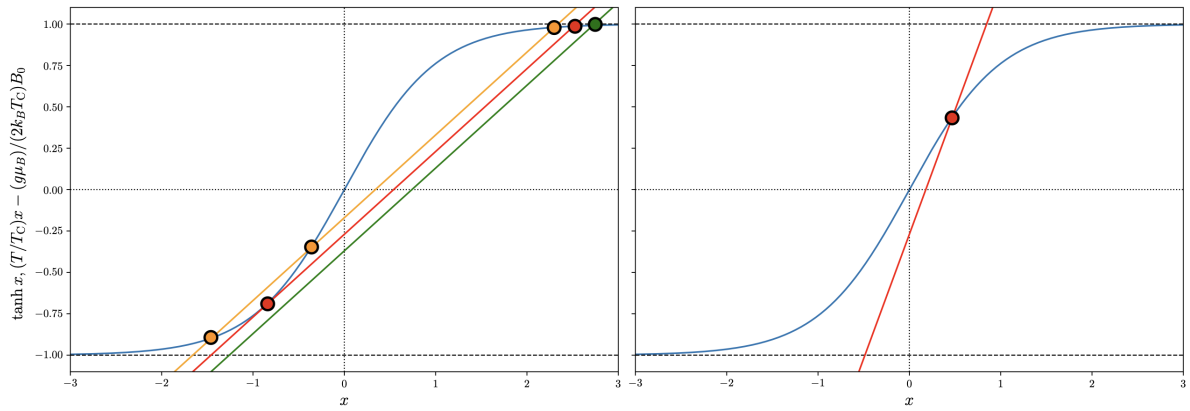



FIGURE 9.12 – Graphical determination of the magnetization in zero applied field, in the case $T < T_\mathrm{C}$ (left, for three different values of the applied field, increasing from orange, to red, to green), and $T > T_\mathrm{C}$ (right).

24. This comes from writing the magnetization equation as

$$\frac{\mathcal{M}}{\mathcal{M}_\infty} = \tanh\left(\frac{g\mu_B B_0}{2k_B T} + \frac{pVK}{2Ng\mu_B k_B T}\mathcal{M}\right) = \tanh\left[\frac{T_\mathrm{C}}{T}\left(\frac{g\mu_B B_0}{2k_B T_\mathrm{C}} + \frac{\mathcal{M}}{\mathcal{M}_\infty}\right)\right] \tag{9.68}$$

### 9.4.5 Critical exponents and comparison to experimental data

We recall that critical exponents appear experimentally when considering properties of the system in the vicinity of the critical temperature, with

$$\mathcal{M}_0 \propto \begin{cases} 0 \\ (T_\mathrm{C} - T)^\beta \end{cases} \qquad C \propto \begin{cases} (T - T_\mathrm{C})^{-\alpha} \\ (T_\mathrm{C} - T)^{-\alpha'} \end{cases} \qquad \chi \propto \begin{cases} (T - T_\mathrm{C})^{-\gamma} & \text{for } T > T_\mathrm{C} \\ (T_\mathrm{C} - T)^{-\gamma'} & \text{for } T < T_\mathrm{C} \end{cases} \tag{9.72}$$

With $\beta \approx 0.34 - 0.38$, $\alpha \approx \alpha' \approx 0.1$, and $\gamma \approx \gamma' \approx 1.3 - 1.4$. The mean-field approximation presented here gives, as we have seen,

$$\beta = \frac{1}{2} \qquad \alpha = \alpha' = 0 \qquad \gamma = \gamma' = 1 \tag{9.73}$$

so these exponents are universal, independent of the number of neighbours $p$ or of the coupling constant $K$. However, they are not in perfect agreement with the experimental values, underlining the shortcomings of the model. Neglecting fluctuations in the mean-field approximation is particularly problematic, since these are expected to become especially large in the vicinity of the critical temperature.

## 9.5 The Ising model of ferromagnetism

### 9.5.1 Introduction

The difficulty of the Heisenberg Hamiltonian lies in the presence of three components of each spin, that do not commute with each other. The **Ising** model does away with this by keeping only the components of the spin along the $Oz$ axis, resulting in a Hamiltonian

$$H_\mathrm{I} = -g\mu_B B_0 \sum_i S_i - K \sum_{\langle i,j \rangle} S_i S_j \tag{9.74}$$

where $S_i = \vec{S}_i \cdot \vec{e}_z$ is by definition the projection of $\vec{S}_i$ on the $Oz$ axis. The commutativity of all the $S_i$ means that the Hamiltonian $H_\mathrm{I}$ may take a diagonal form in an adequate basis. The $x$ and $y$ components are treated in the mean-field approximation, and we try to do better than this in the $z$ direction of the applied field. In fact, there are **exact** solutions of the Ising model, in some cases [25].

### 9.5.2 Solution in the $d = 1$ case

To tackle the Ising model in a one-dimensional setting, we consider that the $N$ spins are placed along a closed loop, which makes all sites equivalent to each other [26]. The Hamiltonian is then

$$H_\mathrm{I} = -\sum_{i=1}^{N} \left( g\mu_B B_0 S_i + K S_i S_{i+1} \right) = -\sum_{i=1}^{N} \left[ \frac{g\mu_B B_0}{2} \left( S_i + S_{i+1} \right) + K S_i S_{i+1} \right] \tag{9.75}$$

with the convention that the site $N+1$ coincides with the site $1$. The system has $2^N$ possible states, since each spin can take either of two values $\pm 1/2$. The partition function is then

$$Z = \sum_{S_1 = -1/2}^{1/2} \cdots \sum_{S_N = -1/2}^{1/2} \exp \left\{ \beta \sum_{i=1}^{N} \left[ \frac{g\mu_B B_0}{2} \left( S_i + S_{i+1} \right) + K S_i S_{i+1} \right] \right\} \tag{9.76}$$

25. In the $d = 1$ case and in the $d = 2$ case provided the applied field is null.
26. Otherwise, edge effects are expected.

We introduce the following function

$$\mathcal{T}(S_i, S_j) = \exp\left\{\beta\left[\frac{g\mu_B B_0}{2}(S_i + S_j) + K S_i S_j\right]\right\} \tag{9.77}$$

so that the partition function takes the forms

$$Z = \sum_{S_1=-1/2}^{1/2} \dots \sum_{S_N=-1/2}^{1/2} \mathcal{T}(S_1, S_2)\,\mathcal{T}(S_2, S_3)\dots\mathcal{T}(S_N, S_1) \tag{9.78}$$

The coefficients $\mathcal{T}(S_i, S_j)$ can be viewed as the elements of a $2\times 2$ matrix, called the **transfer matrix**

$$\mathbb{T} = \begin{bmatrix} \mathcal{T}\left(+\frac{1}{2}, +\frac{1}{2}\right) & \mathcal{T}\left(+\frac{1}{2}, -\frac{1}{2}\right) \\ \mathcal{T}\left(-\frac{1}{2}, +\frac{1}{2}\right) & \mathcal{T}\left(-\frac{1}{2}, -\frac{1}{2}\right) \end{bmatrix} \tag{9.79}$$

From the matrix product relation

$$\sum_{S_k} \mathcal{T}(S_i, S_k)\,\mathcal{T}(S_k, S_j) = \left(\mathbb{T}^2\right)_{S_i, S_j} \tag{9.80}$$

we have, by a chain rule, a simple expression of the partition function

$$Z = \sum_{S_1} \left(\mathbb{T}^N\right)_{S_1, S_1} = \mathrm{Trace}\left(\mathbb{T}^N\right) \tag{9.81}$$

The transfer matrix is real and symmetric, it is therefore possible to find an orthonormal basis in which it is diagonal and takes the form

$$\mathbb{T} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \qquad \text{so that} \qquad \mathbb{T}^N = \begin{bmatrix} \lambda_1^N & 0 \\ 0 & \lambda_2^N \end{bmatrix} \tag{9.82}$$

where $\lambda_1$ and $\lambda_2$ are the (real) eigenvalues of the matrix. Assuming that $|\lambda_2| < |\lambda_1|$, the partition function becomes

$$Z = \lambda_1^N + \lambda_2^N = \lambda_1^N\left(1 + \frac{\lambda_2^N}{\lambda_1^N}\right) \approx \lambda_1^N \tag{9.83}$$

the last equality being in the limit $N \to \infty$. It remains to compute these eigenvalues, which is a straightforward calculation and yields

$$\lambda_1 = \exp\left(\frac{K}{4k_BT}\right)\left\{\cosh\left(\frac{g\mu_B B_0}{2k_BT}\right) + \sqrt{\sinh^2\left(\frac{g\mu_B B_0}{2k_BT}\right) + \exp\left(-\frac{K}{k_BT}\right)}\right\} \tag{9.84}$$

$$\lambda_2 = \exp\left(\frac{K}{4k_BT}\right)\left\{\cosh\left(\frac{g\mu_B B_0}{2k_BT}\right) - \sqrt{\sinh^2\left(\frac{g\mu_B B_0}{2k_BT}\right) + \exp\left(-\frac{K}{k_BT}\right)}\right\} \tag{9.85}$$

From this, we retrieve the Helmholtz free-energy

$$F = -Nk_BT\left\{\frac{K}{4k_BT} + \ln\left[\cosh\left(\frac{g\mu_B B_0}{2k_BT}\right) + \sqrt{\sinh^2\left(\frac{g\mu_B B_0}{2k_BT}\right) + \exp\left(-\frac{K}{k_BT}\right)}\right]\right\} \tag{9.86}$$

which becomes much simpler in the absence of an applied field,

$$F(B_0 = 0, T) = -Nk_BT\ln\left[2\cosh\left(\frac{K}{4k_BT}\right)\right] \tag{9.87}$$

The average magnetization of the system is given by

$$\mathcal{M} = \frac{g\mu_B}{L} \sum_{i=1}^{N} \langle S_i \rangle \tag{9.88}$$

where $L$ is the physical length of the chain. When written out explicitly, we have

$$\mathcal{M} = \frac{1}{L}\frac{1}{Z} \sum_{S_1=-1/2}^{1/2} \dots \sum_{S_N=-1/2}^{1/2} \left( g\mu_B \sum_{i=1}^{N} S_i \right) \exp\left\{ \beta \sum_{i=1}^{N} \left[ g\mu_B B_0 S_i + K S_i S_{i+1} \right] \right\} \tag{9.89}$$

and we recognize the derivative of $Z$ with respect to the applied field,

$$\mathcal{M} = \frac{1}{\beta}\frac{1}{L}\frac{1}{Z}\frac{\partial Z}{\partial B_0} = -\frac{1}{L}\frac{\partial F}{\partial B_0} \tag{9.90}$$

which may be computed explicitly from the expression of $F$ as

$$\mathcal{M} = \frac{N g\mu_B}{2L} \frac{\sinh\left(\dfrac{g\mu_B B_0}{2k_B T}\right)}{\sqrt{\sinh^2\left(\dfrac{g\mu_B B_0}{2k_B T}\right) + \exp\left(-\dfrac{K}{k_B T}\right)}} \tag{9.91}$$

We note that $\mathcal{M} \to 0$ in the limit $B_0 \to 0$, whatever the temperature, so that the magnetic susceptibility stays finite, and **the one-dimensional Ising model presents no phase transition**, contrary to the mean-field approximation, which is therefore blatantly wrong in this $d = 1$ case.

### 9.5.3 Solution in the $d = 2$ case and beyond

The absence of a phase transition in the one-dimensional case is surprising, but the phenomenon exists when considering two-dimensional systems. The computation is however much more complicated : if the $d = 1$ Ising Hamiltonian was solved in 1925, it was not until 1944 that L. Onsager published the first analytical determination of $F$ in the case $d = 2$, restricted to the null-field case, and it took until 1952 for C. N. Yang to give the expression of the spontaneous magnetization $\mathcal{M}_0(T)$ in that case, which gives, for a square lattice in which each site has four neighbours,

$$\mathcal{M}_0(T) = \begin{cases} 0 & \text{for } T > T_{\mathrm{C}} \\ \dfrac{N g\mu_B}{2S}\left[1 - \dfrac{1}{\sinh^4\left(K/2k_B T\right)}\right]^{1/8} & \text{for } T < T_{\mathrm{C}} \end{cases} \tag{9.92}$$

where $S$ is the surface of the crystal. The critical temperature $T_{\mathrm{C}}$ is given by

$$T_{\mathrm{C}} = \frac{K}{2k_B \operatorname{asinh}(1)} \approx 0.567296 \frac{K}{k_B} \tag{9.93}$$

The critical exponents of this model are

$$\beta = \frac{1}{8} \qquad \alpha = \alpha' = 0 \qquad \gamma = \gamma' = \frac{7}{4} \tag{9.94}$$

No analytical solution exists in the $d = 2$ case for $B_0 \neq 0$, nor in the three-dimensional case, but there are elaborate approximate or numerical methods that yield satisfactory results. Among these are :

— Elaborations on the mean-field approach, where small clusters are treated exactly, bathed in a mean-field environment ;
— Numerical methods for a small number of spins, extrapolated to large $N$ ;
— Perturbation methods, where physical quantities are computed as series in a small parameter ;
— Renormalization methods, where the partition function is computed in successive steps, in the vicinity fo the critical point.

# 10

# Heat capacity of solids

In a crystal, atoms are arranged into a regular lattice, but are actually subject to **vibrations** around these equilibrium positions. The energy associated to these vibrations, that has both kinetic and potential parts, is a function of the temperature : the higher the temperature, the larger the vibrational energy. Consequently, there is a **heat capacity** associated to these degrees of freedom, with the following experimental properties :

— At high enough temperatures ($T$ larger than a few hundred K), the heat capacity of solids becomes independent of temperature and takes a value $C = 3Nk_B$, where $N$ is the number of atoms in the solid. This is the **Dulong-Petit law**, that may easily be explained classically from the equipartition theorem (see 3.5.2).

— At low temperatures, the heat capacity goes to zero as $C \propto T^3$. The fact that $C \to 0$ is a quantum property, first elucidated by Einstein in 1907, but the actual scaling with $T$ was explained by Debye in 1912.

The **Einstein and Debye models** of the heat capacity of solids are discussed at length in this chapter.

## 10.1 The Einstein model

### 10.1.1 Presentation

In the Einstein model, the displacements of the $N$ atoms in the lattice with respect to their respective equilibrium positions are considered to be **independent**, that is the effective potential $u_i$ which a given atom $i$ feels is a function of its own displacement $\vec{r}$ but not of the positions of the other atoms. In other words, it is a **mean-field approximation**, of the kind we saw for ferromagnetism in 9.4. For small enough displacements, the potential is [1]

$$u_i(r) = -u_0 + \frac{1}{2}Kr^2 \tag{10.1}$$

where $u_0 > 0$ is the binding energy per atom. Each of these atoms is then viewed as a **three-dimensional harmonic oscillator**, with an energy

$$\varepsilon = \frac{p^2}{2m} - u_0 + \frac{1}{2}Kr^2 \tag{10.2}$$

The crystal itself is then a collection of $N$ independent, three-dimensional harmonic oscillators with identical angular frequencies $\omega_E = \sqrt{K/m}$.

1. This is obtained through a Taylor expansion about the equilibrium position $r = 0$.

### 10.1.2 Partition function

Working in the canonical framework, and since the atoms are both independent and distinguishable (if the separation between them is larger than the amplitudes of the displacements), we have $Z = z^N$, where $z$ is the partition function for a single atom. It is not possible to factorize the three degrees of freedom of a single atom, so these have to be treated simultaneously. If we did this classically, we would find, as per the equipartition theorem, that the mean energy per degree of freedom is $k_B T$ (one half of that due to the kinetic energy, one half to the potential energy), and we would recover the Dulong-Petit law. To explain the actual behaviour at low temperatures, it is mandatory to treat the problem **quantum-mechanically**. We recall that in this case, the energy levels of a three-dimensional harmonic oscillator are given by

$$\varepsilon = -u_0 + \hbar\omega_E \left[ \left( n_x + \frac{1}{2} \right) + \left( n_y + \frac{1}{2} \right) + \left( n_z + \frac{1}{2} \right) \right] \tag{10.3}$$

where $(n_x, n_y, n_z) \in \mathbb{N}^3$. The partition functions $z$ and $Z$ are the computed straightforwardly as

$$z = \exp\left( \frac{u_0}{k_B T} \right) \frac{1}{8 \sinh^3\left( \dfrac{\hbar\omega_E}{2k_B T} \right)} \qquad Z = \exp\left( \frac{N u_0}{k_B T} \right) \left[ \frac{1}{2 \sinh\left( \dfrac{\hbar\omega_E}{2k_B T} \right)} \right]^{3N} \tag{10.4}$$

### 10.1.3 Properties of the system

From this, all properties of the system may be computed, following the general formulas of Chapter 3. These are conveniently expressed introducing the **Einstein temperature** defined as

$$T_E = \frac{\hbar\omega_E}{k_B} \tag{10.5}$$

which depends on the material considered. It decreases for heavier atoms and increases with the stiffness $K$, so with the strength of the binding energy. Using this, we have the Helmholtz free-energy as

$$F = N \left\{ -u_0 + \frac{3}{2} k_B T_E + 3 k_B T \ln\left[ 1 - \exp\left( -\frac{T_E}{T} \right) \right] \right\} \tag{10.6}$$

from which we get the entropy through a derivation with respect to temperature,

$$S = 3N k_B \left\{ -\ln\left[ 1 - \exp\left( -\frac{T_E}{T} \right) \right] + \frac{T_E}{T} \frac{1}{\exp\left( \dfrac{T_E}{T} \right) - 1} \right\} \tag{10.7}$$

and then the average energy $\langle E \rangle = F + TS$, yielding

$$\langle E \rangle = N \left\{ -u_0 + 3 k_B T_E \left[ \frac{1}{2} + \frac{1}{\exp\left( \dfrac{T_E}{T} \right) - 1} \right] \right\} \tag{10.8}$$

where the term $1/2$ comes from the zero-point energy in the quantum treatment of harmonic oscillators. From this we get the **heat capacity** we are after,

$$C = 3N k_B \left( \frac{T_E}{2T} \right)^2 \frac{1}{\sinh^2\left( \dfrac{T_E}{2T} \right)} \tag{10.9}$$

The corresponding curve is shown in Fig. 10.1, along with the limiting behaviours at low and high temperatures. As expected, we recover the Dulong-Petit law in the limit $T \gg T_E$. At low temperatures, we have an exponential decrease to zero heat capacity from the approximate formula

$$C \approx 3Nk_B \left(\frac{T_E}{T}\right)^2 \exp\left(-\frac{T_E}{T}\right) \tag{10.10}$$

Although this finding of $C \to 0$ when $T \to 0$ is a tremendous success of the quantum theory of solids, it does not reproduce the correct experimental behaviour $C \propto T^3$. Fundamentally, this exponential (rather than power-law) decay is due to the existence of a **quantum of excitation** $\hbar\omega_E$ and cannot be overcome unless one finds a way to have a continuous spectrum of energy in the vicinity of the ground state. This will be solved by lifting the hypothesis of independent vibrations from one atom to the other, and treating correctly the **collective behaviour** of the crystal.

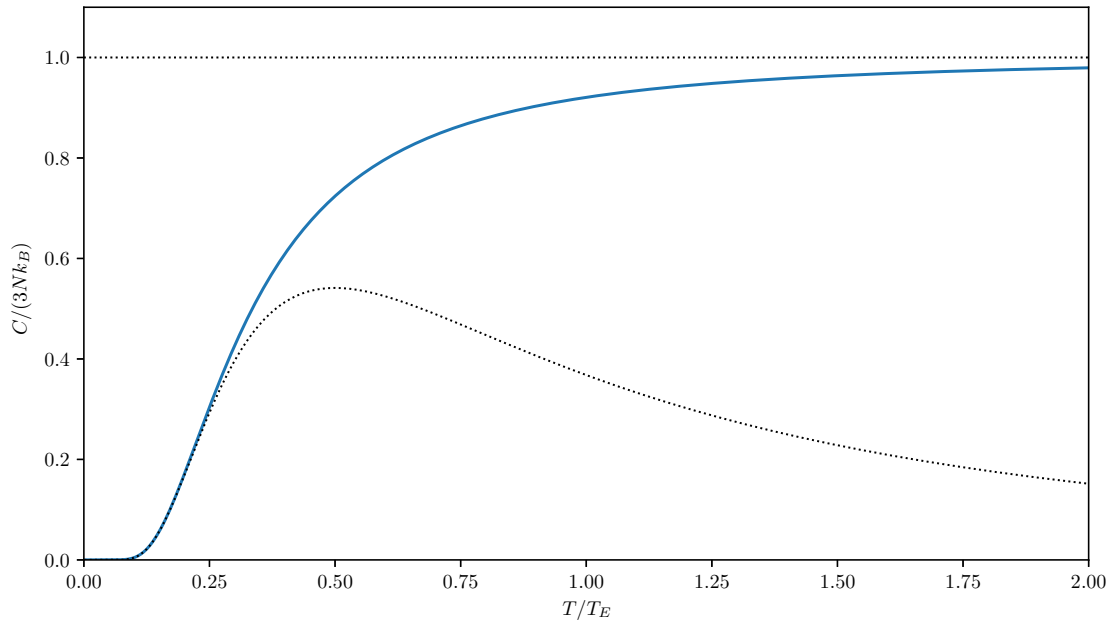



FIGURE 10.1 – The heat capacity of the Einstein model for a three-dimensional monoatomic crystal. Dotted lines mark the expected behaviours at low and high temperatures.

## 10.2 The Debye model

### 10.2.1 Presentation

When an atom is displaced, it necessarily "informs" its neighbours of this displacement, through their mutual interaction. This may be visualized for instance in a one-dimensional chain of identical masses connected by identical springs : if one mass is displaced it will compress one of its adjacent springs, repulsing the neighbouring mass, and dilate the other spring, attracting the other neighbour. Vibrations of the crystal are therefore not independent from one atom to the next but are collective, and propagate as **waves**.

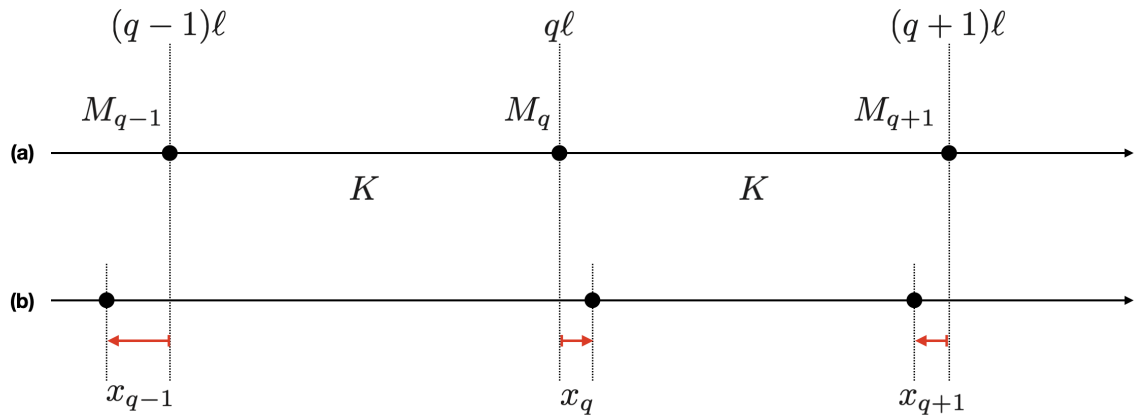



FIGURE 10.2 – One-dimensional monoatomic crystal showing the equilibrium positions of the atoms (a) and their respective displacements (b).

### The one-dimensional crystal

Let us continue with this idea and consider a linear chain of $N$ identical atoms $M_q$ (with $q = 1, \ldots, N$), with masses $m$ and connected through springs of stiffness $K$. To simplify computations, we assume that the chain is closed so that we may consider the atom $M_{N+1}$ to be identical to $M_1$. We note $x_q$ the displacement of $M_q$ with respect to its equilibrium position which is $q\ell = qL/N$, where $L$ is the length of the chain and $\ell = L/N$ is the separation between two atoms at rest. The equation of motion for $M_q$ is easily obtained from the consideration of the elastic forces it is subject to,

$$\frac{\mathrm{d}^2 x_q}{\mathrm{d}t^2} = -\omega_0^2 \left(2x_q - x_{q-1} - x_{q+1}\right) \qquad \omega_0 = \sqrt{\frac{K}{m}} \tag{10.11}$$

These form a system of $N$ coupled differential equations for the $N$ unknowns $x_q$.

### Solutions as progressive waves

Since the system is linear, we seek solutions in terms of longitudinal progressive waves with angular frequency $\omega$ and wavenumber $k$, i.e. , with $x_q = ae^{i(kq\ell - \omega t)}$. Non-trivial solutions of this form are found provided $\omega$ and $k$ obey the **dispersion relation**

$$\omega(k) = 2\omega_0 \left|\sin\left(\frac{k\ell}{2}\right)\right| \tag{10.12}$$

On top of this, the wavenumber $k$ cannot take any value, due to the perdiodic boundary condition $x_{N+1} = x_1$, which leads to

$$k = \frac{2\pi n}{L} \tag{10.13}$$

where $n \in \mathbb{Z}$. Noting also that

$$\omega\left(k + p\frac{2\pi}{\ell}\right) = \omega(k) \tag{10.14}$$

we find that we may restrict ourselves to the **first Brillouin zone** $-\pi/\ell \leqslant k < \pi/\ell$ without losing any solution, and that in this zone, only the $N$ discrete wavenumbers obeying (10.13) - and their associated angular frequencies - are relevant. Note that in a rather large vicinity of the zero wavenumber, we have

a linear dispersion relation $\omega = c|k|$, where $c = \omega_0 \ell$ is the speed of sound in the crystal. The dispersion relation is shown in Fig. 10.3.

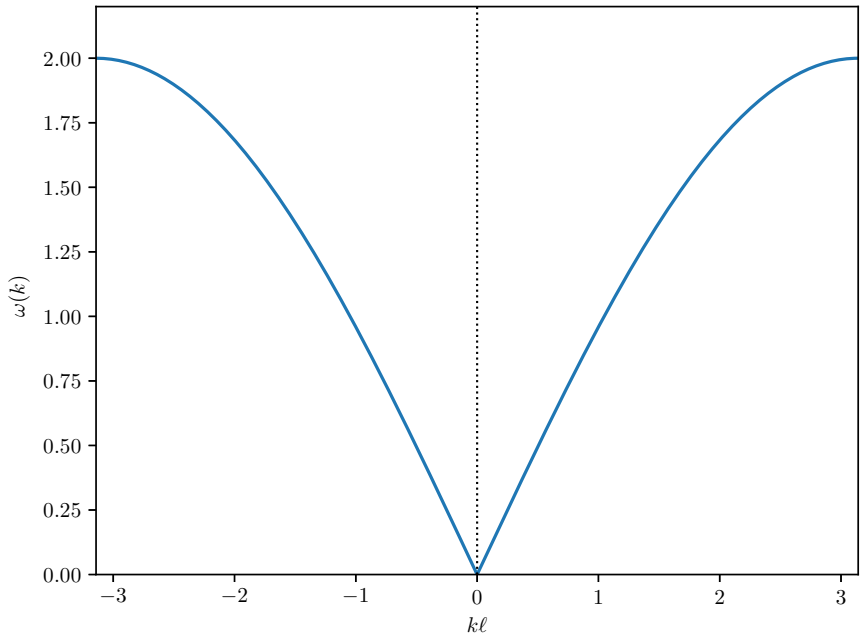



FIGURE 10.3 – The dispersion relation for the one-dimensional monoatomic crystal. The parameters taken here are $K = 1$ and $m = 1$, so that $\omega_0 = 1$ as well.

### Normal modes

Through specific linear combinations of the displacements $x_q$, it is possible to find $N$ new variables that obey differential equations for **independent** harmonic oscillators. These are the **normal modes**

$$\xi_k(t) = \sum_{q=1}^{N} e^{ikq\ell} x_q(t) \tag{10.15}$$

By using the quantization of the wavenumber and the periodicity $x_{N+1} = x_1$, we find indeed that

$$\frac{\partial^2 \xi_k}{\partial t^2} + \omega^2(k)\xi_k = 0 \tag{10.16}$$

so each of the $N$ normal modes evolves as a simple harmonic oscillator with angular frequency $\omega(k)$. A simple example of this is the case of a system of two coupled oscillators, where it is straightforward to exhibit the **symmetric** and **antisymmetric** modes of oscillation.

### 10.2.2 Normal modes and dispersion relation in three dimensions

These ideas generalize to the three-dimensional case, by writing the potential energy up to the second order[2] in the displacements $\vec{r}_i$ of the $N$ atoms as

$$U(\vec{r}_1, \ldots, \vec{r}_N) = -U_0 + \frac{1}{2} \sum_{i,j=1}^{N} \sum_{\alpha,\beta=x,y,z} K_{i\alpha,j\beta} r_{i\alpha} r_{j\beta} \tag{10.17}$$

This energy is that of $3N$ coupled harmonic oscillators, where the diagonal terms are the elastic forces and the non-diagonal ones represent the coupling. In the $3N$-dimensional vector space describing the displacements of the atoms, the matrix whose coefficients are the $K_{i\alpha,j\beta}$ is both real and symmetric, therefore there is a base in which it is diagonal, in other words it is always possible to find a set of linear combinations $\{\rho_{i\alpha}\}$ of the $\{r_{i\alpha}\}$ such that

$$U = -U_0 + \frac{1}{2} \sum_{i=1}^{N} \sum_{\alpha=x,y,z} K'_{i\alpha} \rho_{i\alpha}^2 \tag{10.18}$$

These $\{\rho_{i\alpha}\}$ obey $3N$ equations for independent harmonic oscillators and are therefore the **normal modes** of the problem. To each of these is associated a certain wavevector, which is found to have th efollowing form from the periodic boundary conditions

$$\vec{k} = \frac{p_1}{N_1}\vec{b}_1 + \frac{p_2}{N_2}\vec{b}_2 + \frac{p_3}{N_3}\vec{b}_3 \tag{10.19}$$

where $(p_1, p_2, p_3) \in \mathbb{Z}^3$, $N_1$, $N_2$ and $N_3$ are large numbers of order $N^{1/3}$ such that $N_1 N_2 N_3 = N$, and $\vec{b}_1$, $\vec{b}_2$, $\vec{b}_3$ are the base vectors of the **reciprocal lattice** of the crystal. When restricted to the first Brillouin zone, there are $N$ distinct such wavevectors. On top of that, to each allowed wavevector there are associated three normal modes, that differ by their **polarization** $\lambda = 1, 2, 3$, i.e., the direction in which the displacement occurs when the wave passes, with respect to $\vec{k}$.The **dispersion relation** has three distinct **branches**

$$\omega = \omega_\lambda(\vec{k}) \qquad \lambda = 1, 2, 3 \tag{10.20}$$

For instance, one may have two transverse branches, corresponding to the atoms moving perpendicularly to the direction of the wave, and one longitudinal wave, corresponding to the atoms moving in the direction of the wave. These may have different sound speeds, appearing in the dispersion relations valid at small $|\vec{k}|$,

$$\omega_t = c_t|\vec{k}| \qquad \omega_l = c_l|\vec{k}| \tag{10.21}$$

### 10.2.3 Quantization of the normal modes

#### General expression of crystal properties

The fact that the system may be viewed as $3N$ independent harmonic oscillators (the normal modes) ensures that in the high-temperature limit, the Debye model will recover the Dulong-Petit law. The behaviour at low temperature, however, will differ from the Einstein model because of the **spectrum** of angular frequencies $\omega_\lambda(\vec{k})$ involved, rather than the single $\omega_E$. Quantum-mechanically, the energy of the system is given by the quantum equivalent of Eq. (10.18), i.e.,

$$E_\ell = -U_0 + \sum_{\vec{k},\lambda} \left( n_{\vec{k},\lambda} + \frac{1}{2} \right) \hbar\omega_\lambda(\vec{k}) \qquad |\ell\rangle = \left| \left\{ n_{\vec{k},\lambda} \right\} \right\rangle \tag{10.22}$$

2. The first order is zero by definition of the equilibrium.

where the state $|\ell\rangle$ of the system is specified by the set of occupation numbers [3] of each mode (indexed by their polarization $\lambda$ and their wavevector $\vec{k}$). From this we have the canonical partition function

$$Z = \exp\left(\frac{U_0}{k_B T}\right) \prod_{\vec{k},\lambda} \left( \sum_{n_{\vec{k},\lambda}=0}^{\infty} \exp\left[ -\frac{\hbar\omega_\lambda(\vec{k})}{k_B T} \left( n_{\vec{k},\lambda} + \frac{1}{2} \right) \right] \right) \tag{10.23}$$

The computation is straightforward and yields

$$Z = \exp\left(\frac{U_0}{k_B T}\right) \prod_{\vec{k},\lambda} \frac{1}{2\sinh\left[\dfrac{\hbar\omega_\lambda(\vec{k})}{2k_B T}\right]} \tag{10.24}$$

From this we get the various properties of the crystal, and in particular its average energy

$$\langle E \rangle = -U_0 + \sum_{\vec{k},\lambda} \hbar\omega_\lambda(\vec{k}) \left\{ \frac{1}{2} + \frac{1}{\exp\left[\dfrac{\hbar\omega_\lambda(\vec{k})}{k_B T}\right] - 1} \right\} \tag{10.25}$$

and its heat capacity as the temperature derivative of that average energy

$$C = k_B \sum_{\vec{k},\lambda} \left[\frac{\hbar\omega_\lambda(\vec{k})}{k_B T}\right]^2 \frac{\exp\left[\dfrac{\hbar\omega_\lambda(\vec{k})}{k_B T}\right]}{\left\{\exp\left[\dfrac{\hbar\omega_\lambda(\vec{k})}{k_B T}\right] - 1\right\}^2} \tag{10.26}$$

The allowed wavevectors form a dense lattice in $\vec{k}$-space, allowing for the replacement of the discrete sums by integrals. This is most conveniently done by introducing the **density of normal modes** $\rho(\omega)$ such that the number of normal modes with angular frequencies in $[\omega, \omega + \mathrm{d}\omega]$ is $\rho(\omega)\mathrm{d}\omega$. With this notation, the average energy becomes

$$\langle E \rangle = -\left[ U_0 - \frac{1}{2} \sum_{\vec{k},\lambda} \hbar\omega_\lambda(\vec{k}) \right] + \int_0^{\omega_M} \frac{\hbar\omega\rho(\omega)\mathrm{d}\omega}{\exp\left(\dfrac{\hbar\omega}{k_B T}\right) - 1} = -E_0 + \int_0^{\omega_M} \frac{\hbar\omega\rho(\omega)\mathrm{d}\omega}{\exp\left(\dfrac{\hbar\omega}{k_B T}\right) - 1} \tag{10.27}$$

where $E_0$ is the **dissociation energy** of the crystal, i.e., the energy required to dissociate its constituents at $T = 0$. The upper bound $\omega_M$ comes from the fact that the number of normal modes is finite, with

$$\int_0^{\omega_M} \rho(\omega)\mathrm{d}\omega = 3N \tag{10.28}$$

The density of modes may have a rather complicated shape, depending on the geometry of the crystal. It is usually determined experimentally, and presents two peaks, one near $\omega_M$ corresponding to the longitudinal modes, and one near $\omega_M/2$ or $\omega_M/3$ corresponding to the transverse modes.

### Phonons

The above description is often interpreted in terms of **quasi-particles** called **phonons**. The crystal is then the siege of an ideal gas of these phonons, whose individual states are characterized by a

3. Note that $n_{\vec{k},\lambda} \in \mathbb{N}$

momentum $\vec{p} = \hbar\vec{k}$ and a spin $\lambda$. They have an energy $\varepsilon_\lambda(\vec{p}) = \hbar\omega_\lambda(\vec{k})$ and each of these states may be occupied by a number $n$ of such phonons [4]. Phonons beign indistiguishable, the state of the system is then described by the set of these occupation numbers and its energy is

$$E_\ell = -E_0 + \sum_{\vec{k},\lambda} n_{\vec{k},\lambda}\varepsilon_\lambda(\vec{p}) \tag{10.29}$$

It appears that the average occupation number necessary to recover the expressions below (notably of the average energy) is

$$\langle n(\varepsilon, T)\rangle = \frac{1}{e^{\varepsilon/(k_B T)} - 1} \tag{10.30}$$

implying that the phonons are **bosons** with zero chemical potential, and are therefore not conserved.

### The Debye approximation

At high temperatures, it is seen from Eq. (10.26) that $C \to 3Nk_B$, and we recover the Dulong-Petit law as expected. At low temperatures however, the modes with the lowest angular frequencies will give the largest contributions to the sum. The **Debye approximation** consists in extending to all normal modes the linear dispersion relations (10.21) that are correct at low $\omega$, i.e., $\omega_\lambda(\vec{k}) = c_\lambda|\vec{k}|$. This simple relationship yields the density of normal modes for each polarization in the usual way [5], and from this the total density of normal modes

$$\rho_\lambda(\omega) = \frac{4\pi V\omega^2}{(2\pi c_\lambda)^3} \qquad \rho(\omega) = 2\rho_t(\omega) + \rho_l(\omega) = \frac{3V\omega^2}{2\pi^2\bar{c}^3} \tag{10.31}$$

where we have defined an "average sound speed" $\bar{c}$ through

$$\frac{3}{\bar{c}^3} = \frac{2}{c_t^3} + \frac{1}{c_l^3} \tag{10.32}$$

The maximum angular frequency is in this case the **Debye angular frequency** $\omega_D$ and it is customary to introduce the **Debye temperature** $T_D$, such that [6]

$$\omega_D = \bar{c}\left(\frac{6\pi^2 N}{V}\right)^{1/3} \qquad T_D = \frac{\hbar\omega_D}{k_B} \tag{10.33}$$

With this approximation, the average energy of the system becomes

$$\langle E\rangle = -E_0 + \frac{3V\hbar}{2\pi^2\bar{c}^3}\int_0^{\omega_D} \frac{\omega^3 \mathrm{d}\omega}{\exp\left(\dfrac{\hbar\omega}{k_B T}\right) - 1} = -E_0 + 9Nk_B T\left(\frac{T}{T_D}\right)^3 \int_0^{T_D/T} \frac{x^3 \mathrm{d}x}{e^x - 1} \tag{10.34}$$

The **heat capacity** is simply obtained by deriving this expression with respect to the temperature,

$$C = 9Nk_B\left(\frac{T}{T_D}\right)^3 \int_0^{T_D/T} \frac{x^4 e^x \mathrm{d}x}{(e^x - 1)^2} \tag{10.35}$$

The corresponding curve is shown in Fig. 10.4, along with the low- and high-temperature behaviours discussed in the next paragraph.

4. Note that there is a finite numer of individual states, contrary to what we saw for quantum gases.
5. Recall that each allowed wavevector occupies a volume $(2\pi)^3/V$ in $\vec{k}$-space.
6. The Debye temperature is of order a few hundred K, with extreme cases being boron ($1250\,\mathrm{K}$) and neon ($63\,\mathrm{K}$).

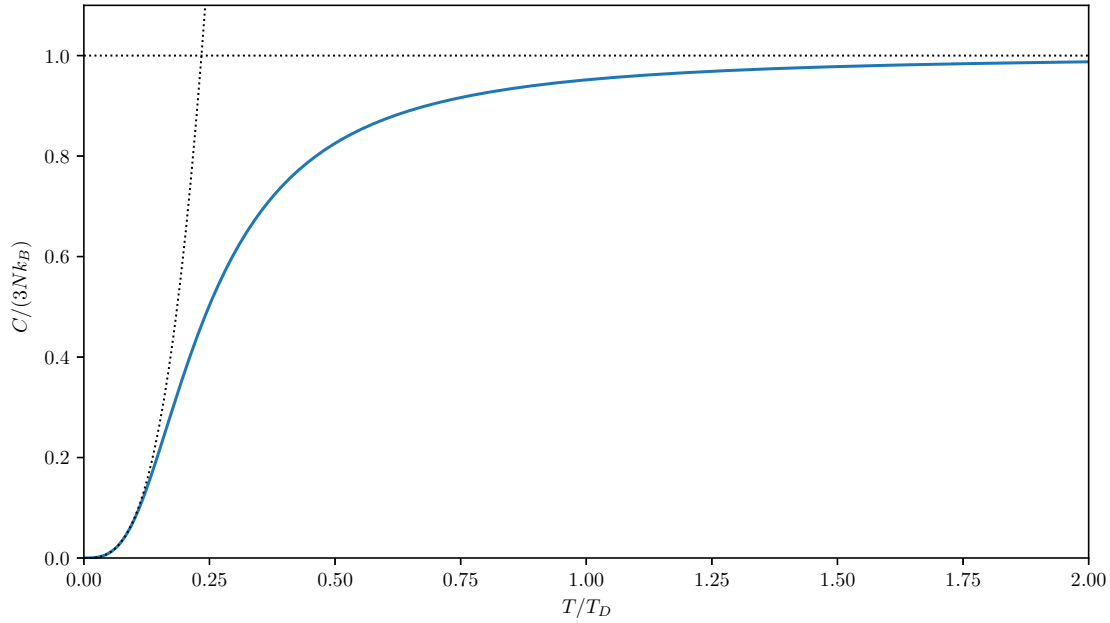



FIGURE 10.4 – The heat capacity of the Debye model for a three-dimensional monoatomic crystal. Dotted lines mark the expected behaviours at low and high temperatures.

#### High and low temperature limits

At high temperature $T \gg T_D$, the upper bound of the integral tends to zero, so that the integrand may be approximated by $x^2$, and the heat capacity becomes

$$C \approx 9Nk_B \left(\frac{T}{T_D}\right)^3 \int_0^{T_D/T} x^2 \mathrm{d}x = 3Nk_B \tag{10.36}$$

as expected to recover the Dulong-Petit law. At low temperatures $T \ll T_D$, the Debye approximation becomes nearly exact, yielding

$$\langle E \rangle \approx -E_0 + \frac{3}{5}\pi^4 Nk_B T \left(\frac{T}{T_D}\right)^3 \tag{10.37}$$

and the heat capacity becomes

$$C \approx \frac{12}{5}\pi^4 Nk_B \left(\frac{T}{T_D}\right)^3 \propto T^3 \tag{10.38}$$

showing that we recover the correct scaling with temperature, contrary to the case of the too simple Einstein model. At intermediate temperatures, the Debye model is an approximation, but it gives good enough results, with the correct behaviour of $C(T)$. Better results may be obtained numerically by integrating over the experimentally-determined density of normal modes.

### 10.2.4 The case of polyatomic crystal lattice

When there is more than one atom per element of the lattice, the results above are slightly altered, but the methods are similar. Taking as an example the case of a one-dimensional crystal with N elements

and two atoms $M_q$ and $M'_q$ per element, as in Fig. 10.5, and assuming that the atoms are identical but are connected by springs of differents stiffnesses $K_1$ and $K_2$, we have

$$m\frac{\mathrm{d}^2 x_q}{\mathrm{d}t^2} = -K_1\left(x_q - x'_q\right) - K_2\left(x_q - x'_{q-1}\right) \tag{10.39}$$

$$m\frac{\mathrm{d}^2 x'_q}{\mathrm{d}t^2} = -K_1\left(x'_q - x_q\right) - K_2\left(x'_q - x_{q+1}\right) \tag{10.40}$$

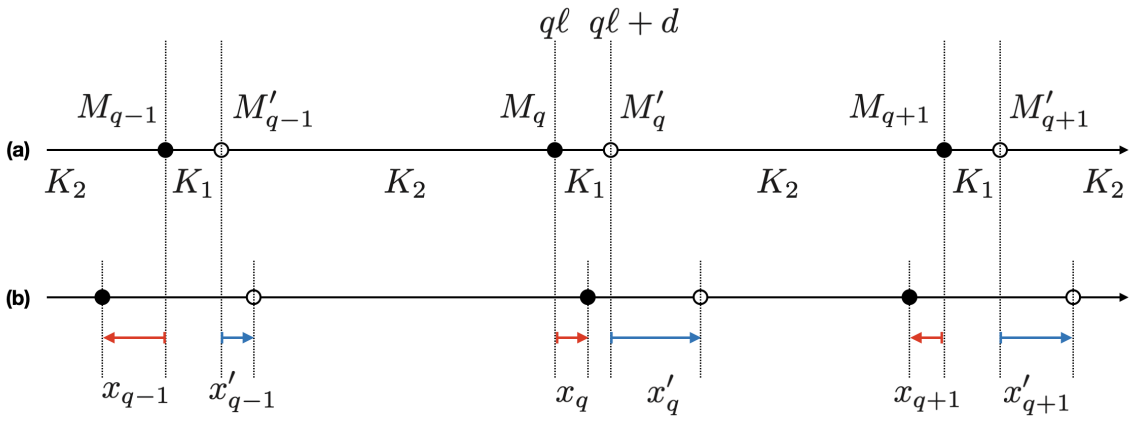



FIGURE 10.5 – One-dimensional diatomic crystal showing the equilibrium positions of the atoms (a) and their respective displacements (b).

By seeking solutions as progressive waves in the form

$$x_q(t) = ae^{i(kq\ell-\omega t)} \qquad x'_q(t) = a'e^{i(kq\ell-\omega t)} \tag{10.41}$$

we find that the existence of non-trivial solutions is only possible if the angular frequency and wave-number obey the following dispersion relation,

$$\omega = \left[\frac{K_1 + K_2 \pm \sqrt{K_1^2 + K_2^2 + 2K_1K_2\cos\left(k\ell\right)}}{m}\right]^{1/2} \tag{10.42}$$

There are thus two branches, depending on the sign in the above equation. These are shown in Fig. 10.6 : the bf acoustic branch is very similar to the one found for the one-dimensional monoatomic crystal (Fig. 10.3), with a linear relation between $k$ and $\omega$ at low $|k|$, while the **optical branch** presents higher, never vanishing angular frequencies,with approximately constant $\omega$ in the vicinity of $k = 0$. In each branch, periodic boundary conditions impose that the wavevector can only take $N$ discrete values. With the two branches, there are thus $2N$ normal modes, equalling the number of atoms in the crystal.

In the more general case of a three-dimensional crystal with $N$ elements and $\nu$ atoms per element, there are $3\nu N$ degrees or freedom, and thus $3\nu N$ normal modes, with $3\nu$ branches and $N$ modes per branch. Among these branches, there are $3$ acousitc branches and $3\nu - 3$ optical branches. To compute the heat capacity associated to the vibrational energy of this crystal, the appropriate method is to use the Debye approximation for the acoustic branches (since $\omega \propto k$) and the Einstein model for the $3\nu - 3$ optical branches (since $\omega$ is approximately constant). This yields [7]

7. We assume the angular frequency of the optical branches to be the same here, for simplicity, and itroduce the corresponding Einstein temperature $T_E = \hbar\omega_{\mathrm{opt}}/k_B$.

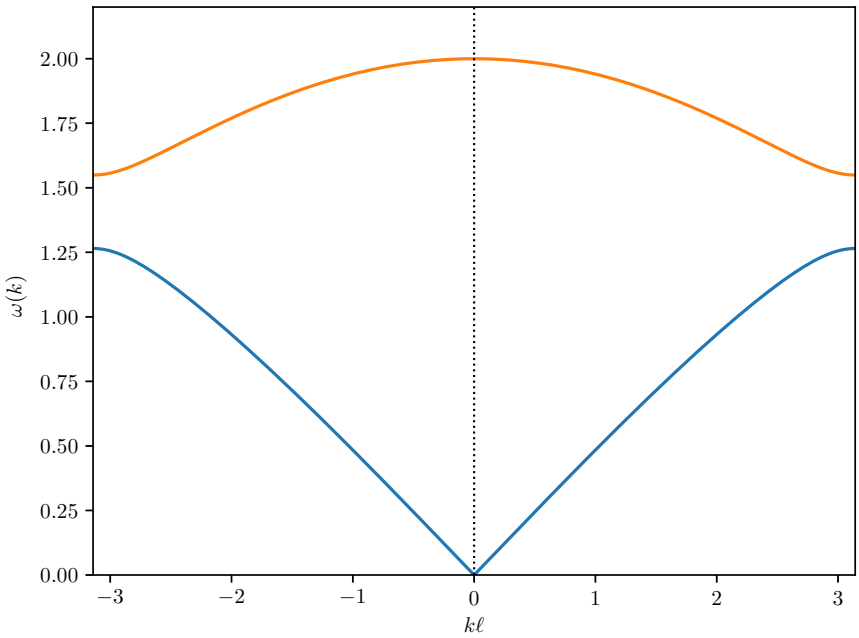



FIGURE 10.6 – The two branches (optical in orange, acoustic in blue) of the dispersion relation for the one-dimensional diatomic crystal. The parameters taken here are $K_1 = 1.2$, $K_2 = 0.8$ and $m = 1$.

$$C = 3Nk_B \left\{ 3\left(\frac{T}{T_D}\right)^3 \int_0^{T_D/T} \frac{x^4 e^x \mathrm{d}x}{(e^x-1)^2} + (\nu-1)\left(\frac{T_E}{2T}\right)^2 \frac{1}{\sinh^2\left(\frac{T_E}{2T}\right)} \right\} \tag{10.43}$$

# Part II

# Out-of-equilibrium systems

# 11

# Elementary laws of transport phenomena

So far, we have only discussed properties of systems at equilibrium. However, we mentioned in several instances how a system evolved when an internal constraint was relaxed, a topic we will discuss at length in chapter 13. With this chapter, we begin the discussion of properties of **out-of-equilibrium** statistical physics and the general laws that govern **transport phenomena** in near-equilibrium. The computation of the **transport coefficients** that will be introduced here will be one of the topics of **kinetic theory**, treated in chapter 15.

## 11.1 General considerations

### 11.1.1 The problem

The questions raised by considering out-of-equilibrium statistical physics are the following : *Is there a general principle allowing to describe the irreversible evolution of a system initially out of equilibrium ? What happens in a system that is maintained out of equilibrium ?*

Equilibrium thermodynamics only provides general principles for the final equilibrium state of the system, and there is no such general principle for out-of-equilibrium system unless one of the two following conditions is met :

— The evolution is slow, so that the system may be considered to be **locally at equilibrium** ;

— The system is "not far" from equilibrium, in which case a **linear approach** may be undertaken.

This may be illustrated by Fig. 11.1, where three systems, each in one of the experimental conditions discussed in Chapters 2 to 4, evolve from out-of-equilibrium states, provoked for instance by the relaxation of an internal constraint, towards equilibrium states. The hypotheses expressed above are represented by the fact that the out-of-equilibrium states are "close" to the states extremizing the respective thermodynamical potentials, and that the motion of the representative points towards the equilibria should be slow.

Failing those conditions, there is no general principle and each system should be treated with its particularities.

### 11.1.2 Local equilibrium

Consider an experiment in which a single drop of ink is placed within a volume of water. We may ask how the density of "ink molecules" $n(\vec{r}, t)$ evolves in time and space. Since the system consisting of the ink and water may be considered isolated, the final state of the system will be that maximizing the entropy, which in this case means that $n$ shall eventually be uniform. How does that come about ?

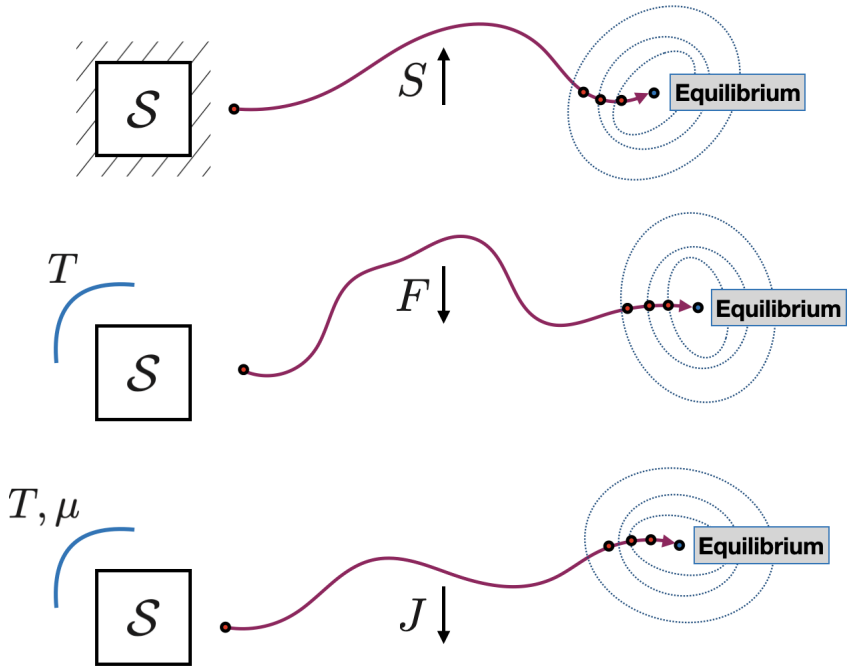



FIGURE 11.1 – Evolution of initially out-of-equilibrium systems towards their respective equilibria that extremize their respective thermodynamical potentials when an internal constraint is relaxed. The dotted closed lines represent isocontours of the thermodynamical potentials and the red dots represent local equilibrium states on the way to a global equilibrium.

Within a small volume $\mathrm{d}v$ around position $\vec{r}$, there are $\mathrm{d}N = n(\vec{r},t)\mathrm{d}v$ ink molecules, which will undergo random collisions among themselves but also with water molecules. This will tend to uniformize the density $n$ within that small volume, over a time-scale $\tau_{\mathrm{eq}}$ that is of the order of a few collisional times [1] $\tau_{\mathrm{coll}}$. Now, the overall evolution of $n$ within the full volume is characterized by another time-scale $\tau_{\mathrm{ev}}$. The hypothesis of **local equilibrium** holds provided that

$$\tau_{\mathrm{eq}} \ll \tau_{\mathrm{ev}} \tag{11.1}$$

Under that condition, we may consider that at each instant $t$ (measured with a precision $\delta t$ significantly larger than $\tau_{\mathrm{eq}}$ but also significantly smaller than $\tau_{\mathrm{ev}}$), the molecules within $\mathrm{d}v$ are at equilibrium among themselves. We may then safely define that density $n(\vec{r},t)$ of ink molecules, but also the **energy density** $e(\vec{r},t)$, the **entropy density** $s(\vec{r},t)$ (that is the volumetric entropy of a system with particle density $n$ and energy density $e$ at this place and time), the **local temperature** $T(\vec{r},t)$ and the **local chemical potential** $\mu(\vec{r},t)$. The latter two are defined through

$$\frac{1}{T} = \frac{\partial s}{\partial e} \qquad\qquad \frac{\mu}{T} = -\frac{\partial s}{\partial n} \tag{11.2}$$

Any measurement apparatus will average these over timescales $\delta t$ such that $\tau_{\mathrm{eq}} \ll \delta t \ll \tau_{\mathrm{ev}}$, resulting in smooth evolution curves such as shown in Fig. 11.2.

### 11.1.3 Conservation laws

Within the small volume $\mathrm{d}v$, local equilibrium does not change the number of particles $\mathrm{d}N$ or the energy $\mathrm{d}E$, since (elastic) collisions conserve both of these. However, these same conservation laws

1. That is the typical time between two collisions.

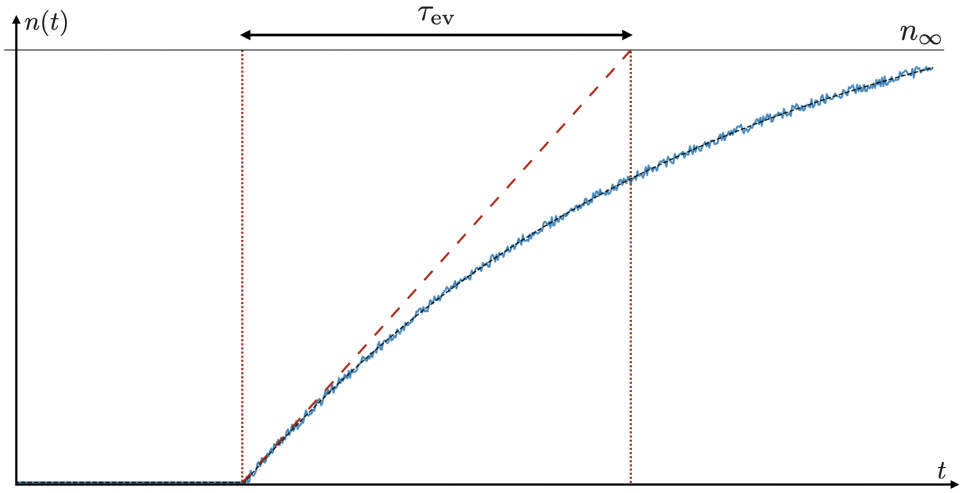



FIGURE 11.2 – Time evolution of the density of particles. The evolution at small time-scales is shown in blue, and the evolution observed when averaging over a time-scale $\delta\tau$ such that $\tau_{\mathrm{eq}} \ll \delta\tau \ll \tau_{\mathrm{ev}}$ is shown in black.

mean that the final, global equilibrium where $n$ and $e$ are uniform throughout can only be attained by a slow **transport** of particles and energy in the whole volume. To each conserved quantity there corresponds a certain **transport phenomenon** :

— **Particle diffusion** is associated to the conservation of particle number
— **Heat diffusion** is associated to the conservation of energy
— **Viscosity** is associated to the conservation of momentum

It should be noted that the timescales for these different phenomena may be quite different in some instances. If the transport of one quantity is much faster than that of another, the former may be considered at equilibrium while we treat the transport of the latter.

## 11.2 Particle diffusion : Fick's law

### 11.2.1 Position of the problem

Let us consider an isothermal system ($T$ uniform) and assume that it is also an isobaric system ($P$ uniform) so that global motions are prevented. We consider the evolution of the density $n(\vec{r},t)$ of a certain type of particles, which incidentally implies that there are at least two types of particles [2]. Under local equilibrium, the chemical potential is $\mu(\vec{r},t) = \mu[n(\vec{r},t),T,P]$ and the global equilibrium is attained when that is uniform throughout (see the discussion in 8.1.3), so that $n$ is also uniform.

### 11.2.2 Particle density current

For a small surface $\mathrm{d}\vec{S} = \mathrm{d}S\vec{u}$ within the volume, where $\vec{u}$ is a unit vector perpendicular to the surface, we write $\mathrm{d}^2N$ to stand for the number of particles crossing $\mathrm{d}S$ in the duration $\mathrm{d}t$ (counted positively if in the direction of $\vec{u}$, negatively otherwise). For elementary $\mathrm{d}S$ and $\mathrm{d}t$, that number is

2. Indeed, for a simple fluid, the equation of state $f(T,P,n) = 0$ with uniform $T$ and $P$ implies that $n$ is also uniform. We therefore must have at least two types of particles, with densities $n_1$ and $n_2$, and the equation of state $f(T,P,n_1,n_2) = 0$ with $T$ and $P$ uniform no longer implies the uniformity of the density of particles. We note that to treat the problem of **self-diffusion**, it is necessary to tag some particles, e.g. with a radioactive isotope, so we do effectively have two types of particles here as well.

proportional to $\mathrm{d}S\mathrm{d}t$ and it may be written introducing the **particle density current** $\vec{J}_N$ as

$$\mathrm{d}^2N = \vec{J}_N \cdot \mathrm{d}\vec{S}\mathrm{d}t \tag{11.3}$$

The scalar product makes sense because the current of particles must be a vector quantity, and if we imagine a situation in which this current is parallel to the surface $\mathrm{d}S$, there is no net flux of particles from one side of the surface to the other, so $\mathrm{d}^2N = 0$.

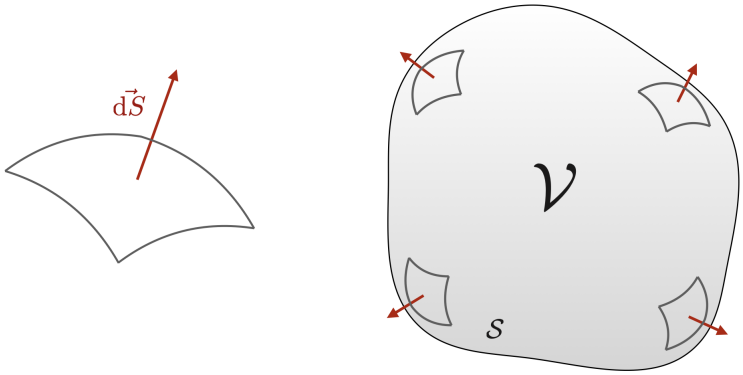



FIGURE 11.3 – Particle density current and the conservation of particles

The particle density current may be related to the local density $n$ and the local fluid velocity (averaged over a mesoscopic volume) $\vec{v}$ as

$$\vec{J}_N = n\vec{v} \tag{11.4}$$

This can be obtained by considering that particles crossing the surface during $\mathrm{d}t$ were in the cylinder with base $\mathrm{d}S$ and height $v\mathrm{d}t\cos\theta$, where $\theta$ is the angle between $\vec{v}$ and the normal to the surface element. Indeed, with a fluid velocity $\vec{v}$, the number of particles crossing $\mathrm{d}S$ in the duration $\mathrm{d}t$ is

$$\mathrm{d}^2N = n \times v\mathrm{d}t\cos\theta \times \mathrm{d}S = n\vec{v} \cdot \mathrm{d}\vec{S}\mathrm{d}t \tag{11.5}$$

### 11.2.3 Conservation of particles

Now consider a fixed volume $\mathcal{V}$, bounded by a surface $\mathcal{S}$ through which particles may either enter or leave the volume. The number of particles within that volume is, by definition,

$$N(t) = \iiint_{\mathcal{V}} n(\vec{r},t)\mathrm{d}v \tag{11.6}$$

so the variation of that number, given that the volume is fixed, is

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \iiint_{\mathcal{V}} \frac{\partial n}{\partial t}\mathrm{d}v \tag{11.7}$$

but since the particles are conserved, that variation is entirely due to particles entering and leaving the volume through its boundary, so making use of the particle density current we also have

$$\frac{\mathrm{d}N}{\mathrm{d}t} = -\iint_{\mathcal{S}} \vec{J}_N \cdot \mathrm{d}\vec{S} \tag{11.8}$$

where the minus sign is because the vectors $\mathrm{d}\vec{S}$ are by convention oriented outwards, as in the figure. Equating those two expressions and making use of the Green-Ostrogradski divergence theorem, we have

$$\iiint_{\mathcal{V}} \frac{\partial n}{\partial t} \mathrm{d}v + \iint_{\mathcal{S}} \vec{J}_N \cdot \mathrm{d}\vec{S} = \iiint_{\mathcal{V}} \left( \frac{\partial n}{\partial t} + \vec{\nabla} \cdot \vec{J}_N \right) \mathrm{d}v = 0 \tag{11.9}$$

Since this must hold for any volume $\mathcal{V}$, we conclude that the conservation of particles is expressed, in differential terms, as the following exact relation

$$\frac{\partial n}{\partial t} + \vec{\nabla} \cdot \vec{J}_N = 0 \tag{11.10}$$

### 11.2.4 Law of (linear) transport

At equilibrium, the chemical potential is uniform and there is no transport of particles, which is expressed mathematically as $\vec{J}_N = \vec{0}$. Out of equilibrium, $\mu$ is no longer uniform, so its gradient is non-zero, $\vec{\nabla}\mu \neq \vec{0}$, and the particle current density is also non-zero $\vec{J}_N \neq \vec{0}$. If the system is not too far from equilibrium, $||\vec{\nabla}\mu||$ is small and we may expect that there is a **linear response** of $\vec{J}_N$ to $\vec{\nabla}\mu$

$$\vec{J}_N = -\alpha \vec{\nabla} \mu \tag{11.11}$$

where $\alpha > 0$ to comply with the discussion of 8.1.3 that particles flow from regions with high chemical potential to regions with lower chemical potential. This is a phenomenological approximation, that needs to be validated through experiments. Given that $\mu$ is a function of the particle density $n$, it is more common to express this relation in terms of the gradient of $n$, resulting in **Fick's law**

$$\vec{J}_N = -D \vec{\nabla} n \qquad\qquad D = \alpha \left( \frac{\partial \mu}{\partial n} \right)_{T,P} \tag{11.12}$$

The coefficient $D$, which depends on $n$, $T$, $P$, and on the composition of the ambient medium, is called the **diffusion coefficient**. We note that this discussion assumes that the medium is homogeneous, isothermal, and isobaric, that local equilibrium applies, and that fluctuations of $n$ are small enough for the linear approximation to hold.

The dimension of the diffusion coefficient is $\mathrm{m}^2\,\mathrm{s}^{-1}$, but it is more usually expressed in $\mathrm{cm}^2\,\mathrm{s}^{-1}$. Typical values in gases are of the order $\sim 0.1 - 1\,\mathrm{cm}^2\,\mathrm{s}^{-1}$, while it is much smaller in liquids $\sim 10^{-5}\,\mathrm{cm}^2\,\mathrm{s}^{-1}$, and even smaller in solids, with for instance $\sim 10^{-10}\,\mathrm{cm}^2\,\mathrm{s}^{-1}$ for helium gas in silica $\mathrm{SiO}_2$, and $\sim 10^{-30}\,\mathrm{cm}^2\,\mathrm{s}^{-1}$ for aluminium within copper.

### 11.2.5 Diffusion equation

The conservation law and Fick's law may now be combined, using the definition of the **Laplacian operator** $\Delta$ to obtain an equation for $n$ that constitutes the **diffusion equation**

$$\frac{\partial n}{\partial t} = D \Delta n \tag{11.13}$$

Note that this assumes $D$ to be homogeneous, otherwise we default to the more general relation

$$\frac{\partial n}{\partial t} = \vec{\nabla} \cdot \left( D \vec{\nabla} n \right) \tag{11.14}$$

This is a partial derivative equation, that requires boundary and initial conditions to be given in order to be solved, i.e. we need to know $n(\vec{r},0)$ for all $\vec{r}$ in the volume, and $n(\vec{r}_0,t)$ for all $t$ and all positions $\vec{r}_0$ belonging to the boundaries of the system.

### 11.2.6 Application : evolution of a one-dimensional solution

Working with a one-dimensional system, the diffusion equation becomes

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} \tag{11.15}$$

We take a known, not necessarily simple, initial condition and take the system to be infinite in both directions, so the particle density and its gradient are null at infinity,

$$n(x,0) = n_0(x) \qquad n(x\to\pm\infty,t) = 0 \qquad \frac{\partial n}{\partial x}(x\to\pm\infty,t) = 0 \tag{11.16}$$

In fact, the gradient should go to zero fast [3]. We check that these conditions are compatible with the conservation of the number of particles as

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \int_{-\infty}^{+\infty}\frac{\partial n}{\partial t}\mathrm{d}x = D\int_{-\infty}^{+\infty}\frac{\partial^2 n}{\partial x^2}\mathrm{d}x = \left[\frac{\partial n}{\partial x}\right]_{-\infty}^{+\infty} = 0 \tag{11.17}$$

The average position of the particles is given by

$$\langle x\rangle(t) = \frac{1}{N}\int_{-\infty}^{+\infty} x n(x,t)\mathrm{d}x \tag{11.18}$$

and it is straightforward to show that this average position is constant in time. Indeed, we have

$$\frac{\mathrm{d}\langle x\rangle}{\mathrm{d}t} = \frac{1}{N}\int_{-\infty}^{+\infty} x\frac{\partial n}{\partial t}\mathrm{d}x = \frac{D}{N}\int_{-\infty}^{+\infty} x\frac{\partial^2 n}{\partial x^2}\mathrm{d}x = \frac{D}{N}\left\{\left[x\frac{\partial n}{\partial x}\right]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty}\frac{\partial n}{\partial x}\mathrm{d}x\right\} \tag{11.19}$$

The fully integrated term is zero because of the fast decrease of the gradient at infinity, and the other term is also null

$$\frac{\mathrm{d}\langle x\rangle}{\mathrm{d}t} = \frac{D}{N}\left[n\right]_{-\infty}^{+\infty} = 0 \tag{11.20}$$

We may therefore choose this average position to be the origin of the coordinates system, $\langle x\rangle(t) = 0$. Now the spread of the particle distribution in space will be characterized by the variance

$$\langle x^2\rangle(t) = \frac{1}{N}\int_{-\infty}^{+\infty} x^2 n(x,t)\mathrm{d}x \tag{11.21}$$

and it is similarly straightforward to show that this increases linearly with time

$$\frac{\mathrm{d}\langle x^2\rangle}{\mathrm{d}t} = 2D \tag{11.22}$$

Indeed, we have

$$\frac{\mathrm{d}\langle x^2\rangle}{\mathrm{d}t} = \frac{1}{N}\int_{-\infty}^{+\infty} x^2\frac{\partial n}{\partial t}\mathrm{d}x = \frac{D}{N}\int_{-\infty}^{+\infty} x^2\frac{\partial^2 n}{\partial x^2}\mathrm{d}x = \frac{D}{N}\left\{\left[x^2\frac{\partial n}{\partial x}\right]_{-\infty}^{+\infty} - 2\int_{-\infty}^{+\infty} x\frac{\partial n}{\partial x}\mathrm{d}x\right\} \tag{11.23}$$

3. The condition for $N$ to be finite is that $n$ should go to zero faster than $1/x$, which corresponds to the gradient going to zero faster than $1/x^2$, but in order for the variance $\langle x^2\rangle$ to be meaningful, the density $n$ itself should go to zero faster than $1/x^3$, corresponding to the gradient going to zero faster than $1/x^4$.

so that taking into account the fast enough decrease of the gradient at infinity, we have

$$\frac{\mathrm{d}\langle x^2\rangle}{\mathrm{d}t} = -2\frac{D}{N}\int_{-\infty}^{+\infty} x\frac{\partial n}{\partial x}\mathrm{d}x = -2\frac{D}{N}\left\{[xn]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} n\mathrm{d}x\right\} = 2D \tag{11.24}$$

So if for instance the initial distribution $n_0(x)$ is a Dirac delta function $n_0(x) = \delta(x)$, for which the standard deviation is zero, the spread of the distribution of particles in space goes as

$$\sigma_x(t) = \sqrt{\langle x^2\rangle(t)} = \sqrt{2Dt} \tag{11.25}$$

In fact, in this case the distribution $n(x,t)$ may be computed exactly, and it takes a Gaussian form

$$n(x,t) = \frac{N}{\sqrt{4\pi Dt}}\exp\left(-\frac{x^2}{4Dt}\right) \tag{11.26}$$

The diffusion equation being linear, its solution for any initial distribution $n_0(x)$ is found through a convolution of this initial distribution by that above Gaussian response. In other words, the Gaussian is **Green's function** for the one-dimensional diffusion operator.

To get an idea of orders of magnitude, the time required for diffusion to spread an initially point-like distribution over a scale $\Delta x$ is

$$t = \frac{(\Delta x)^2}{2D} \tag{11.27}$$

so for instance, in a gas with $D \sim 0.1\,\mathrm{cm^2\,s^{-1}}$ we find $t \sim 5\,\mathrm{s}$ for $\Delta x = 1\,\mathrm{cm}$, but $t > 10\,\mathrm{h}$ for $\Delta x = 1\,\mathrm{m}$. Diffusion is really slow ! Yet if someone opens a bottle of perfume at the other end of the room, the scent will reach us much faster than this, because other processes are much more efficient at transporting molecules, in this case convection. Another famous example of the slowness of diffusion is the experiment set-up in 1872 by Kelvin, in which a copper sulfate solution was made to diffuse in water within a 5 m-glass tube. The solution is not yet homogeneous today [4], which is not surprising since the caracteristic time for $\Delta x = 5\,\mathrm{m}$ and $D \sim 10^{-5}\,\mathrm{cm^2\,s^{-1}}$ is about 340 years...

## 11.3 Electric conductivity : Ohm's law

### 11.3.1 Response to an electric field

Consider a system of electrons, subject to an external electric field $\vec{\mathcal{E}} = -\vec{\nabla}V$. As we saw earlier in 8.2.2, and in particular equation (8.24), the chemical potential in this situation is

$$\mu(\vec{r}) = \mu_0\left[T, n(\vec{r})\right] - eV(\vec{r}) \tag{11.28}$$

If the system is at equilibrium, $\mu$ is uniform, but if the electric current is allowed to flow (in a circuit for instance), then $\mu$ is no longer uniform even though $n$ may be. A current of electrons sets in. The particle density current is then

$$\vec{J}_N = -\alpha\vec{\nabla}\mu = \alpha e\vec{\nabla}V = -\alpha e\vec{\mathcal{E}} \tag{11.29}$$

From this, we define the **electric current density** $\vec{j} = -ne\vec{v} = -e\vec{J}_N$, leading to **Ohm's law**

$$\vec{j} = \sigma\vec{\mathcal{E}} \qquad\qquad \sigma = e^2\alpha \tag{11.30}$$

where $\sigma$ is the **electric conductivity**.

4. https://cen.acs.org/articles/91/i33/Old-Experiment-Even-Older-Curiosities.html

### 11.3.2 Einstein's relation for semi-conductors

The $\alpha$ parameter appears in both definitions of the diffusion coefficient $D$ and of the electric conductivity. This suggests that the two are related, which can be shown as

$$\frac{D}{\sigma} = \frac{\alpha \left(\frac{\partial \mu}{\partial n}\right)_T}{e^2 \alpha} = \frac{1}{e^2}\left(\frac{\partial \mu}{\partial n}\right)_T = \frac{1}{e^2}\left(\frac{\partial \mu_0}{\partial n}\right)_T \tag{11.31}$$

However, in general this relationship is rather written in terms of the **mobility** $\gamma$ of the electrons, that is defined from $\vec{v} = \gamma \vec{\mathcal{E}}$. Since $\vec{j} = -ne\vec{v} = -ne\gamma\vec{\mathcal{E}}$, we have $\sigma = -ne\gamma$ and so

$$\frac{D}{\gamma} = -\frac{n}{e}\left(\frac{\partial \mu_0}{\partial n}\right)_T \tag{11.32}$$

For a semiconductor, we recall that equation (8.34) states that $\mu_0 = -k_B T \ln[\alpha(T)] + k_B T \ln n$ and so we end up with **Einstein's relation** linking the diffusion coefficient for the electrons and their mobility

$$\frac{D}{\gamma} = -\frac{k_B T}{e} \tag{11.33}$$

### 11.3.3 External field and particle density imbalance

We saw previously that in the absence of an external electric field $\vec{\mathcal{E}} = \vec{0}$ but in the presence of particle density imbalance $\vec{\nabla} n \neq \vec{0}$, a particle density current appeared with

$$\vec{J_N} = -D\vec{\nabla} n \tag{11.34}$$

Conversely, we just saw that in the presence of an external electric field $\vec{\mathcal{E}} \neq \vec{0}$ but in the absence of particle density imbalance $\vec{\nabla} n = \vec{0}$, an electric current $\vec{j}$ appeared, that could be associated to a particle density current with

$$\vec{J_N} = -\frac{1}{e}\vec{j} = -\frac{\sigma}{e}\vec{\mathcal{E}} \tag{11.35}$$

When both sources of imbalance are present, and provided we remain in the small deviations case where the linear regime is applicable, we must conclude that a particle density current appears, with

$$\vec{J_N} = -D\vec{\nabla} n - \frac{\sigma}{e}\vec{\mathcal{E}} \tag{11.36}$$

until, at equilibrium, this current density vanishes, leading to sustained, non-zero electric field and particle density gradient with

$$\vec{\nabla} n = -\frac{\sigma}{eD}\vec{\mathcal{E}} = \frac{\sigma}{eD}\vec{\nabla} V = -\frac{n\gamma}{D}\vec{\nabla} V \tag{11.37}$$

This may be integrated to show that the particle density in a semi-conductor adapts to the electric potential $V$ in order to obey the Boltzmann statistics

$$n = n_0 \exp\left(\frac{eV}{k_B T}\right) = n_0 \exp\left(-\frac{E}{k_B T}\right) \tag{11.38}$$

where $E = qV = -eV$ is the energy of the charge $q = -e$ in the potential $V$.

## 11.4 Thermal conductivity : Fourier's law

### 11.4.1 Different types of energy transfer

In a system where the temperature $T$ is not uniform, an energy transfer is required to recover equilibrium. This transfer may be through **radiation**, through **convection**, a natural or forced transfer via macroscopic motions of the medium that is difficult to eliminate, or through **conduction**, an energy transfer that does not involve a transfer of matter.

### 11.4.2 Heat conduction

As we did previously, we assume that local equilibrium holds so we may speak of the local temperature $T(\vec{r},t)$. Given a small surface $\mathrm{d}\vec{S} = \mathrm{d}S\vec{u}$, we consider the heat $\delta^2 Q$ crossing this surface between $t$ and $t+\mathrm{d}t$, counted positively if the heat transfer is in the direction of $\vec{u}$. This heat is proportional to $\mathrm{d}S$ and $\mathrm{d}t$, leading to the definition of the **heat flux** $\vec{J}_Q$ through

$$\delta^2 Q = \vec{J}_Q \cdot \mathrm{d}\vec{S}\mathrm{d}t \tag{11.39}$$

We note that this quantity has the dimensions of a power per unit surface. As with the particle density current, in the case of equilibrium, i.e. for $\vec{\nabla}T = \vec{0}$, we expect this heat flux $\vec{J}_Q$ to be zero, and to be different from zero if the system is out of equilibrium $\vec{\nabla}T \neq \vec{0}$. Therefore, in a linear regime of small deviations, the two vectors should be related by the phenomenological **Fourier's law**

$$\vec{J}_Q = -K\vec{\nabla}T \tag{11.40}$$

where $K$ is the **thermal conductivity**, whose dimensions are $\mathrm{W\,m^{-1}\,K^{-1}}$. For solids, typical orders of magnitude are $K \sim 0.1 - 10\,\mathrm{W\,m^{-1}\,K^{-1}}$ in non-metallic solids, $K \sim 100\,\mathrm{W\,m^{-1}\,K^{-1}}$ in metals. For liquids, we have $K \sim 0.5\,\mathrm{W\,m^{-1}\,K^{-1}}$ in water, but $K \sim 80\,\mathrm{W\,m^{-1}\,K^{-1}}$ in liquid sodium. For gases, we have typically $K \sim 10^{-3} - 10^{-1}\,\mathrm{W\,m^{-1}\,K^{-1}}$.

### 11.4.3 Energy conservation

To obtain an equation describing the evolution of the temperature field $T(\vec{r},t)$, we require a supplementary equation relating $\vec{J}_Q$ and $T$, expressing the conservation of energy. Assume that the evolution occurs at constant pressure, and consider a small element of mass $\delta m = \rho\delta V$ that is fixed so we can ignore its bulk kinetic energy. In the time interval $\mathrm{d}t$, the energy of this element varies according to the first principle of thermodynamics

$$\mathrm{d}\left(\delta U\right) = \delta W + \delta Q = -P\mathrm{d}\left(\delta V\right) + \delta Q \tag{11.41}$$

which can be rewritten by introducing the enthalpy $\delta H$ of the system

$$\mathrm{d}\left(\delta U + P\delta V\right) = \mathrm{d}\left(\delta H\right) = \delta Q \tag{11.42}$$

By dividing by the time interval $\mathrm{d}t$, the heat flux appears

$$\frac{\mathrm{d}\delta H}{\mathrm{d}t} = \frac{\delta Q}{\mathrm{d}t} = -\vec{\nabla}\cdot\vec{J}_Q\delta V \tag{11.43}$$

where the last equality arises from applying the Green-Ostrogradski theorem to the rate of heat exchange across the boundaries of an elementary volume, much like what was done for the conservation of the total number of particles earlier. Indeed, we have

$$\frac{\delta Q}{\mathrm{d}t} = -\iint \vec{J}_Q\cdot\mathrm{d}\vec{S} = -\iiint \vec{\nabla}\cdot\vec{J}_Q\mathrm{d}v = -\vec{\nabla}\cdot\vec{J}_Q\delta V \tag{11.44}$$

Now the left-hand side of the above equation may be written using the **heat capacity at constant pressure per unit mass** $c_p$ since

$$\frac{\mathrm{d}\delta H}{\mathrm{d}t} = \frac{\partial \delta H}{\partial T}\frac{\partial T}{\partial t} = \rho c_p \delta V \frac{\partial T}{\partial t} \tag{11.45}$$

leading to the following second relation between the heat flux and the temperature

$$\rho c_p \frac{\partial T}{\partial t} + \vec{\nabla}\cdot\vec{J_Q} = 0 \tag{11.46}$$

Note that energy is always conserved, but that temperature may vary even in an isolated system, for instance because of chemical reactions or radioactive decay. In that case, the enthalpy per unit volume $h(T,x)$ will be a function not only of the temperature but also of the advancement parameter $x$ of the reaction, and the left-hand side of (11.43) above will now be

$$\frac{1}{\delta V}\frac{\mathrm{d}\delta H}{\mathrm{d}t} = \left(\frac{\partial h}{\partial T}\right)_{P,x}\frac{\partial T}{\partial t} + \left(\frac{\partial h}{\partial x}\right)_{T,P}\frac{\mathrm{d}x}{\mathrm{d}t} = \rho c_p \frac{\partial T}{\partial t} + \left(\frac{\partial h}{\partial x}\right)_{T,P}\frac{\mathrm{d}x}{\mathrm{d}t} \tag{11.47}$$

### 11.4.4 Heat equation

Combining these two equations relating the temperature and the heat flux, and assuming that the thermal conductivity is uniform, we obtain the **heat equation**

$$\frac{\partial T}{\partial t} = D_{\mathrm{th}}\Delta T \qquad\qquad D_{\mathrm{th}} = \frac{K}{\rho c_p} \tag{11.48}$$

where $D_{\mathrm{th}}$ is the **thermal diffusivity**, which has the same dimensions as the particle diffusion coefficient $D$, i.e., $\mathrm{m}^2\,\mathrm{s}^{-1}$. If $K$ is not uniform, the heat equation defaults to the more general form

$$\rho c_p \frac{\partial T}{\partial t} = \vec{\nabla}\cdot\left(K\vec{\nabla}T\right) \tag{11.49}$$

An example of an application of the heat equation is a thermal shock, where a medium initially at thermal equilibrium with $T(x,t) = T_0$ for $t < 0$ is perturbed at $t > 0$ by imposing that $T(0,t) = T_1 \neq T_0$. Dimensional analysis suggests to look for similarity solutions in the form

$$T(x,t) = f\left(\frac{x}{\sqrt{D_{\mathrm{th}}t}}\right) \tag{11.50}$$

Other examples include the daily and annual variations of the temperature of the ground as a function of depth, or the problem of thermal contact between two systems at temperatures $T_1$ and $T_2$.

### 11.4.5 Wiedemann-Franz law

In a metal, heat and charge transport are both due to free electrons, so there must be a relationship between the thermal conductivity $K$ and the electric conductivity $\sigma$. Indeed, such a relationship exists, and is called the **Wiedemann-Franz law**

$$\frac{K}{\sigma} = \frac{\pi^2}{3}\left(\frac{k_B}{e}\right)^2 T \tag{11.51}$$

# 12

# General principles of near-equilibrium thermodynamics

In this chapter, we outline **Onsager's theory** for near-equilibrium thermodynamics. This approach, dating from the 1930s, is justified if the transport phenomena involved are slow enough that local equilibrium holds, and if the deviations from equilibrium are small, so that a linear treatment is possible. This is a general, well-structured formulation leading to new and non-trivial physics as we shall see.

## 12.1 Irreversible evolution of a continuous system

### 12.1.1 Entropy density variations

Consider a system at local equilibrium, with particle density $n(\vec{r},t)$, energy density $e(\vec{r},t)$, and entropy density $s(\vec{r},t)$. The local temperature $T(\vec{r},t)$ and chemical potential $\mu(\vec{r},t)$ are defined from

$$\frac{1}{T}=\frac{\partial s}{\partial e} \qquad\qquad \frac{\mu}{T}=-\frac{\partial s}{\partial n} \tag{12.1}$$

This should be viewed as stemming from the fact that the dependence on space and time of $s$ is actually through $n$ and $e$, i.e., since from $S=S(E,N,V)$ we have $s=s[e(\vec{r},t),n(\vec{r},t)]$. More generally, the entropy will be a function of a number of extensive variables $X_i$ (not counting the volume), so that

$$S=S\left(\{X_i\},V\right) \tag{12.2}$$

from which we get the dependency of the entropy density

$$s=s\left[\{x_i(\vec{r},t)\}\right] \tag{12.3}$$

where $x_i$ is the volumetric density of $X_i$ (i.e., the amount of $X_i$ per unit volume). The variations of the entropy are given by

$$\mathrm{d}S=\sum_i Y_i \mathrm{d}X_i+\frac{P}{T}\mathrm{d}V \tag{12.4}$$

where we introduce the **conjugate intensive variable** $Y_i$ of $X_i$ in $S$, defined by

$$Y_i=\left(\frac{\partial S}{\partial X_i}\right)_{X_j,V} \tag{12.5}$$

Examples of conjugate intensive variables are $Y_e=1/T$ and $Y_n=-\mu/T$. Working with a given fixed volume, and moving to volumetric quantities, this translates to a variation of the entropy density as

$$\mathrm{d}s=\sum_i Y_i \mathrm{d}x_i \tag{12.6}$$

### 12.1.2 Currents and conservation laws

For any extensive variable $X_i$, we may define a **current density** $\vec{J}_i$ associated to it, by considering a small surface $\mathrm{d}\vec{\Sigma} = \mathrm{d}\Sigma\vec{u}$ and writing that the amount of $X_i$ crossing this surface between $t$ and $t+\mathrm{d}t$, counted algebraically, is proportional to $\mathrm{d}\Sigma$ and $\mathrm{d}t$, with

$$\mathrm{d}^2 X_i = \vec{J}_i \cdot \mathrm{d}\vec{\Sigma}\mathrm{d}t \tag{12.7}$$

Locally, energy and particle number are conserved. If the same is true with the variable $X_i$, we may derive local equations expressing these various conservation laws [1]

$$\frac{\partial n}{\partial t} + \vec{\nabla}\cdot\vec{J}_n = 0 \qquad \frac{\partial e}{\partial t} + \vec{\nabla}\cdot\vec{J}_e = 0 \qquad \frac{\partial x_i}{\partial t} + \vec{\nabla}\cdot\vec{J}_i = 0 \tag{12.8}$$

### 12.1.3 Entropy current and production

Since the elementary variation of entropy density is related to those of the densities $x_i$ by

$$\mathrm{d}s = \sum_i Y_i \mathrm{d}x_i = \frac{1}{T}\mathrm{d}e - \frac{\mu}{T}\mathrm{d}n + \dots \tag{12.9}$$

we may analogously define an **entropy density current** through

$$\vec{J}_s = \sum_i Y_i \vec{J}_i = \frac{1}{T}\vec{J}_e - \frac{\mu}{T}\vec{J}_n + \dots \tag{12.10}$$

such that the scalar product $\vec{J}_s \cdot \mathrm{d}\vec{\Sigma}$ represents the (algebraic) amount of entropy crossing the surface $\mathrm{d}\Sigma$ per unit time. Note that this vector is closely related to the heat flux, defined in Chapter 11,

$$\vec{J}_s = \frac{1}{T}\vec{J}_Q \tag{12.11}$$

Considering a volume $\mathcal{V}$ bounded by a fixed surface $\partial\mathcal{V}$, the variation of the entropy within this volume between $t$ and $t+\mathrm{d}t$ is given by the second principle of thermodynamics for an irreversible process

$$\mathrm{d}S > \frac{\delta Q}{T} \tag{12.12}$$

where $\delta Q$ is the heat gained by the system in the process, and $T$ its temperature. The two sides of this inequality can be expressed in terms of the entropy density and entropy density current, respectively [2]

$$\mathrm{d}S = \mathrm{d}t\int_{\mathcal{V}} \frac{\partial s}{\partial t}\mathrm{d}V > \frac{\delta Q}{T} = -\int_{\mathcal{S}} \frac{1}{T}\vec{J}_Q \cdot \mathrm{d}\vec{\Sigma}\mathrm{d}t = -\int_{\mathcal{V}} \vec{\nabla}\cdot\vec{J}_s \mathrm{d}V\mathrm{d}t \tag{12.13}$$

Since $\mathrm{d}t > 0$, this translates into a local equation for the entropy density

$$\frac{\partial s}{\partial t} + \vec{\nabla}\cdot\vec{J}_s = \dot{s} > 0 \tag{12.14}$$

where $\dot{s}$ thus defined is the **local rate of entropy density production** through irreversibility of the process under study.

1. We already saw the first of these, Eq. (11.10). Note that $\vec{J}_e$ is not equal to the heat flux $\vec{J}_Q$, as we shall see later.
2. We recall that the minus sign stems from the outwards orientation of the surface vectors $\mathrm{d}\vec{\Sigma}$.

### 12.1.4 Thermodynamic forces

This rate of entropy density production may be expressed as a function of the currents $\vec{J}_i$ and of quantities known as thermodynamic forces. To see this, we start with

$$\frac{\partial s}{\partial t} = \sum_i Y_i \frac{\partial x_i}{\partial t} \qquad \vec{J}_s = \sum_i Y_i \vec{J}_i \tag{12.15}$$

combining these in the equation giving $\dot{s}$ above, we have

$$\dot{s} = \sum_i \left[ Y_i \frac{\partial x_i}{\partial t} + \vec{\nabla} \cdot \left( Y_i \vec{J}_i \right) \right] \tag{12.16}$$

Now the divergence terms may be expanded using the relation $\vec{\nabla} \cdot \left( f\vec{A} \right) = f\vec{\nabla} \cdot \vec{A} + \vec{A} \cdot \vec{\nabla} f$, yielding

$$\dot{s} = \sum_i \left[ Y_i \left( \frac{\partial x_i}{\partial t} + \vec{\nabla} \cdot \vec{J}_i \right) + \vec{J}_i \cdot \vec{\nabla} Y_i \right] = \sum_i \vec{J}_i \cdot \vec{\nabla} Y_i \tag{12.17}$$

using the conservation equations in the last step. Introducing the **thermodynamic forces** $\vec{\mathcal{F}}_i = \vec{\nabla} Y_i$, we have the local entropy density production rate as

$$\dot{s} = \sum_i \vec{J}_i \cdot \vec{\mathcal{F}}_i \tag{12.18}$$

For instance, in the simple case where only particle number density and energy density determine the entropy density we have

$$\dot{s} = \vec{J}_e \cdot \vec{\nabla} \left( \frac{1}{T} \right) - \vec{J}_n \cdot \vec{\nabla} \left( \frac{\mu}{T} \right) \tag{12.19}$$

and thus we have the associated thermodynamic forces

$$\vec{\mathcal{F}}_e = \vec{\nabla} \left( \frac{1}{T} \right) \qquad \vec{\mathcal{F}}_n = -\vec{\nabla} \left( \frac{\mu}{T} \right) \tag{12.20}$$

### 12.1.5 Linear regime and transport coefficients

If all thermodynamic forces are null, $\vec{\mathcal{F}}_i = \vec{0}$, the system is at equilibrium. For instance, in the simple case $s(e, n)$ above, that would mean that $T$ and $\mu$ are uniform. Consequently, all currents $\vec{J}_i$ are also null. Conversely, if any of those thermodynamic forces is non-zero, the system is out of equilibrium, and currents $\vec{J}_i$ set in to recover an equilibrium situation. If those thermodynamic forces are small enough, a **linear regime** may be expected to hold between them and the currents, i.e. we postulate that

$$\vec{J}_i = \sum_j L_{ij} \vec{\mathcal{F}}_j \tag{12.21}$$

where $\mathbf{L} = [L_{ij}]$ is a matrix whose elements are the **phenomenological Onsager coefficients** $L_{ij}$. It is clear that Fick's, Ohm's, and Fourier's laws are particular cases of such a situation. To determine these coefficients, microscopic models from kinetic theory (Chapter 15) and experiments are necessary. The Onsager framework cannot determine them explicitly, but it can provide constraints on their values.

### 12.1.6 Properties of $[L_{ij}]$

The positivity of $\dot{s}$ implies several properties of the $[L_{ij}]$ matrix. Indeed, we have, irrespective of the specific values of the thermodynamic forces

$$\dot{s} = \sum_{ij} L_{ij} \vec{\mathcal{F}}_i \cdot \vec{\mathcal{F}}_j > 0 \tag{12.22}$$

so $L_{ij}$ is a representation of a positive definite quadratic form, and we can use particular cases to obtain properties of the coefficients. For instance, if all but one thermodynamic force $\vec{\mathcal{F}}_i$ is null, we get $L_{ii} > 0$, and therefore the matrix diagonal is strictly positive

$$\forall i \quad L_{ii} > 0 \tag{12.23}$$

Also, if only two of those thermodynamic forces are non-zero, and we assume that they are parallel to $\vec{e}_x$, we have, writing $X = |\vec{\mathcal{F}}_i|/|\vec{\mathcal{F}}_j|$,

$$L_{ii} X^2 + (L_{ij} + L_{ji}) X + L_{jj} > 0 \tag{12.24}$$

which shows that the discriminant of this second-order polynomial should be negative

$$(L_{ij} + L_{ji})^2 - 4 L_{ii} L_{jj} < 0 \tag{12.25}$$

Finally, it may be shown that in the absence of a magnetic field [3], the matrix is symmetric $L_{ij} = L_{ji}$. We therefore have the **Onsager relations** relating the various phenomenological coefficients

$$L_{ii} > 0 \qquad L_{ij} = L_{ji} \qquad L_{ii} L_{jj} > L_{ij} L_{ji} = L_{ij}^2 \tag{12.26}$$

## 12.2 Application to properties of electrical conductors

Consider an electrical conductor subject to gradients of temperature $\vec{\nabla} T$ and of potential $\vec{\nabla} V$, in which electrons are free to circulate. In stationary state, these gradients are constant in time, as is the density $n$ of electrons. We assume that there is no magnetic field, and that the conductor is isotropic.

### 12.2.1 Currents and thermodynamic forces

As already seen in Chapter 8, the local chemical potential of electrons is

$$\mu(\vec{r}) = \mu_0 \left[ T(\vec{r}), n \right] - eV(\vec{r}) \tag{12.27}$$

and as we saw above the inhomogeneity of $\mu$ or $T$ leads to non-zero particle density and energy currents $\vec{J}_n$ and $\vec{J}_e$. As written earlier, the rate of entropy density creation is then

$$\dot{s} = \vec{J}_e \cdot \vec{\nabla} \left( \frac{1}{T} \right) - \vec{J}_n \cdot \vec{\nabla} \left( \frac{\mu}{T} \right) \tag{12.28}$$

However, in practice we measure not $\vec{J}_e$ but the heat flux

$$\vec{J}_Q = T \vec{J}_s = \vec{J}_e - \mu \vec{J}_n \tag{12.29}$$

3. When a magnetic field is present, $L_{ij}(\vec{B}) = L_{ji}(-\vec{B})$

where this expression comes from subtracting the energy carried by particles from the energy current $\vec{J}_e$. From the above relations, we can obtain the expression of $\dot{s}$ for the alternate set of currents $(\vec{J}_n, \vec{J}_Q)$

$$\dot{s} = \left(\vec{J}_Q + \mu\vec{J}_n\right).\vec{\nabla}\left(\frac{1}{T}\right) - \vec{J}_n.\vec{\nabla}\left(\frac{\mu}{T}\right) = \vec{J}_Q.\vec{\nabla}\left(\frac{1}{T}\right) - \frac{1}{T}\vec{J}_n.\vec{\nabla}\mu \tag{12.30}$$

So the thermodynamic forces associated to these currents are [4]

$$\vec{\mathcal{F}}_n = -\frac{1}{T}\vec{\nabla}\mu \qquad \vec{\mathcal{F}}_Q = \vec{\nabla}\left(\frac{1}{T}\right) \tag{12.31}$$

and the linear regime relations between these and the currents are written as

$$\vec{J}_n = -\frac{L_{11}}{T}\vec{\nabla}\mu + L_{12}\vec{\nabla}\left(\frac{1}{T}\right) \qquad \vec{J}_Q = -\frac{L_{21}}{T}\vec{\nabla}\mu + L_{22}\vec{\nabla}\left(\frac{1}{T}\right) \tag{12.32}$$

### 12.2.2 Thermoelectric effects

#### Isothermal electric conductivity

We consider a case where there is no gradient of temperature or density but there is a non-zero electric potential gradient, in which case the above relations simplify to

$$\vec{J}_n = -\frac{L_{11}}{T}\vec{\nabla}\mu \qquad \vec{J}_Q = -\frac{L_{21}}{T}\vec{\nabla}\mu \tag{12.33}$$

With a uniform temperature and density, the gradient of $\mu$ reduces to $\vec{\nabla}\mu = -e\vec{\nabla}V = e\vec{\mathcal{E}}$, where $vec\mathcal{E}$ is the electric field, and the electric current density is $\vec{j} = -e\vec{J}_n$ so we recover **Ohm's law** with an expression of the electric conductivity $\sigma$

$$\vec{j} = \sigma\vec{\mathcal{E}} \qquad \sigma = \frac{L_{11}e^2}{T} \tag{12.34}$$

#### Thermal conductivity

Now consider a situation in which energy is transported without matter transport. We then have $\vec{J}_n = \vec{0}$ but $\vec{\nabla}T \neq \vec{0}$. The relations above now yield

$$\vec{0} = -\frac{L_{11}}{T}\vec{\nabla}\mu + L_{12}\vec{\nabla}\left(\frac{1}{T}\right) \qquad \vec{J}_Q = -\frac{L_{21}}{T}\vec{\nabla}\mu + L_{22}\vec{\nabla}\left(\frac{1}{T}\right) \tag{12.35}$$

The first of these allows to relate the chemical potential gradient to that of temperature

$$\vec{\nabla}\mu = -\frac{L_{12}}{TL_{11}}\vec{\nabla}T \tag{12.36}$$

which may be reinjected into the second equation to obtain **Fourier's law** and an expression of the thermal conductivity $K$ as a function of the phenomenological coefficients $L_{ij}$. Indeed,

$$\vec{J}_Q = -\frac{L_{21}}{T} \times \left(-\frac{L_{12}}{TL_{11}}\right)\vec{\nabla}T - \frac{L_{22}}{T^2}\vec{\nabla}T \tag{12.37}$$

4. $\vec{\mathcal{F}}_n$ is not the same as previously written in Eq. (12.20), because we changed the sets of currents from $(\vec{J}_e, \vec{J}_n)$ to $(\vec{J}_Q, \vec{J}_n)$.

which is easily rearranged into

$$\vec{J}_Q = -K\vec{\nabla}T \qquad K = \frac{L_{22}L_{11} - L_{12}L_{21}}{T^2 L_{11}} \tag{12.38}$$

with the Onsager relations ensuring then that $K > 0$, as it should.

### Seebeck effect (1826)

The relation (12.36) above shows that in an open circuit (so that there is no circulation of electrons), a thermal gradient leads to the emergence of a chemical potential gradient. This is the **Seebeck effect**. We rewrite this relation as

$$\frac{1}{e}\vec{\nabla}\mu = \varepsilon(T)\vec{\nabla}T \tag{12.39}$$

introducing the **Seebeck coefficient**, also called **thermopower**, which depends on the material,

$$\varepsilon(T) = -\frac{L_{12}}{eTL_{11}} \tag{12.40}$$

Typical values of thermopowers for various metals or alloys are of the order of $\sim 10\,\mu\mathrm{V}\cdot\mathrm{K}^{-1}$.

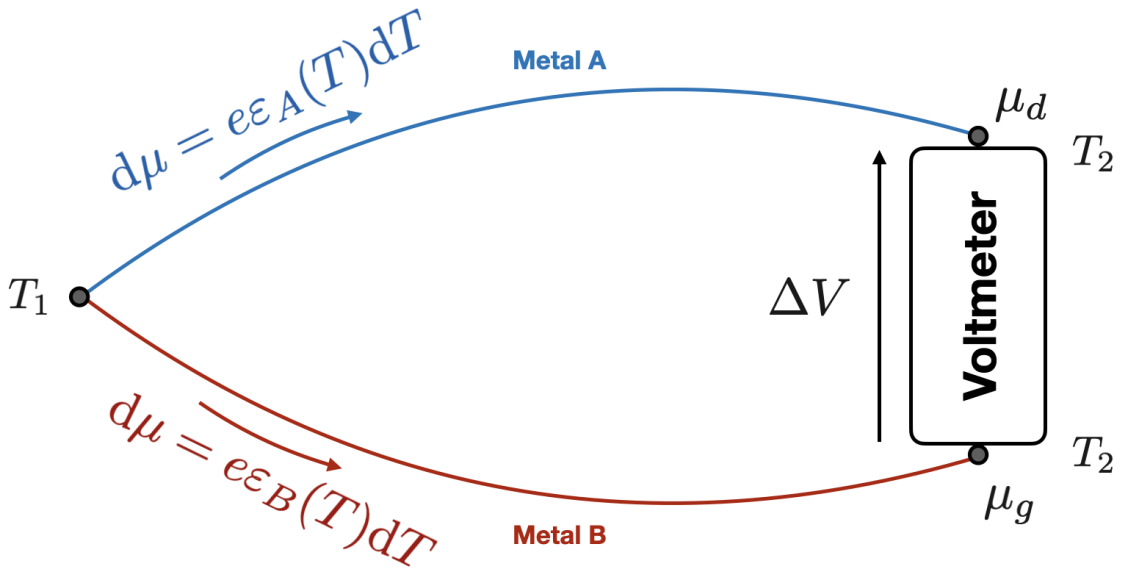



FIGURE 12.1 – Principle of a thermocouple

The fact that a temperature variation implies a change in chemical potential and that this change may be different for different materials is the basis of the temperature-measuring devices known as **thermocouples**, as shown in Fig. 12.1. Take two wires made of different metals, $A$ and $B$, and put them in contact as indicated in the figure. At the other end, insert a voltmeter. The two junctions are at different temperatures $T_1$ and $T_2$, thereby imposing a temperature gradient (that is the same in both metals), thus also a chemical potential gradient (that is different in both metals). The voltmeter lets no electric current through, as it has a very large electric resistance, but it offers no thermal resistance, so both ends of it are at $T_2$. Along each metal wire we have $\mathrm{d}\mu = e\varepsilon_X(T)\mathrm{d}T$ with $X = A$ or $X = B$,

and so the chemical potential difference between the two ends of the voltmeter is

$$\mu_d - \mu_g = \int_{T_1}^{T_2} e\varepsilon_A(T)\mathrm{d}T + \int_{T_2}^{T_1} e\varepsilon_B(T)\mathrm{d}T = \int_{T_1}^{T_2} e\left[\varepsilon_A(T) - \varepsilon_B(T)\right]\mathrm{d}T \tag{12.41}$$

Now, we recall that $\mu = \mu_0\left(T, n, P\right) - eV$, so we may write this in terms of electric potential difference. It should be noted that the density of charges may be (and indeed is) different at both ends of the voltmeter ($n_d \neq n_g$), but in fact, a voltmeter is actually measuring the difference of chemical potential expressed in terms of an electric potential, i.e.,

$$\Delta V = V_d - V_g = -\frac{\Delta\mu}{e} = \frac{\mu_g - \mu_d}{e} \tag{12.42}$$

so that we have

$$\Delta V = \int_{T_1}^{T_2} \left(\varepsilon_B - \varepsilon_A\right)\mathrm{d}T \tag{12.43}$$

If we know $T_2$, then this relation gives $T_1$ through the measurement of the electric potential difference $\Delta V$ [5].

We note that if we know the electric conductivity, the thermal conductivity and the thermopower

$$\sigma = \frac{L_{11}e^2}{T} \qquad K = \frac{L_{22}L_{11} - L_{12}L_{21}}{T^2 L_{11}} \qquad \varepsilon = -\frac{L_{12}}{eTL_{11}} \tag{12.44}$$

all elements of the $\mathbf{L}$ matrix are known.

**Interpretation of the thermopower**

A physical interpretation of the thermopower arises from eliminating the gradient of $\mu$ between $\vec{J}_n$ and $\vec{J}_Q$, using Eq. (12.32). Indeed, we have

$$\frac{\vec{J}_n}{L_{11}} = -\frac{1}{T}\vec{\nabla}\mu + \frac{L_{12}}{L_{11}}\vec{\nabla}\left(\frac{1}{T}\right) \qquad \frac{\vec{J}_Q}{L_{21}} = -\frac{1}{T}\vec{\nabla}\mu + \frac{L_{22}}{L_{21}}\vec{\nabla}\left(\frac{1}{T}\right) \tag{12.45}$$

which we can subtract from one another, yielding

$$\frac{\vec{J}_Q}{L_{21}} - \frac{\vec{J}_n}{L_{11}} = \frac{L_{22}}{L_{21}}\vec{\nabla}\left(\frac{1}{T}\right) - \frac{L_{12}}{L_{11}}\vec{\nabla}\left(\frac{1}{T}\right) \tag{12.46}$$

so that

$$\vec{J}_Q = \frac{L_{21}}{L_{11}}\vec{J}_n + \left(L_{22} - \frac{L_{21}L_{12}}{L_{11}}\right)\vec{\nabla}\left(\frac{1}{T}\right) \tag{12.47}$$

In the first term, we recognize, using the symmetry property $L_{21} = L_{12}$, that

$$\frac{L_{21}}{L_{11}} = -e\varepsilon T \tag{12.48}$$

and in the second, we recognize the thermal conductivity, so that eventually

$$\vec{J}_Q = -K\vec{\nabla}T - e\varepsilon T\vec{J}_n \tag{12.49}$$

Thus, the heat flux consists of two terms, the first one associated to heat transport by **diffusion**, and the second one to heat transport by **convection**. The expression of the entropy density current

$$\vec{J}_s = -\frac{K}{T}\vec{\nabla}T - e\varepsilon\vec{J}_n \tag{12.50}$$

shows that $-e\varepsilon$ is the **entropy transported by each electron**.

5. In practice, $T_2$ is measured by another device, or we have two thermocouples, one of them plunging into a bath of known temperature.

### Peltier effect (1834)

The **Peltier effect** is the emergence of a heat flux induced by an electric current, under uniform temperature conditions. From $\vec{\nabla} T = \vec{0}$ we indeed have, from equation (12.49), that

$$\vec{J}_Q = -e\varepsilon T \vec{J}_n = \varepsilon T \vec{j} = \Pi \vec{j} \tag{12.51}$$

where $\Pi = \varepsilon T$ is the **Peltier coefficient**. This effect may be observed at the junction of two conductors $A$ and $B$ maintained at a uniform $T$ through contact with a thermostat, as shown in Fig. 12.2.

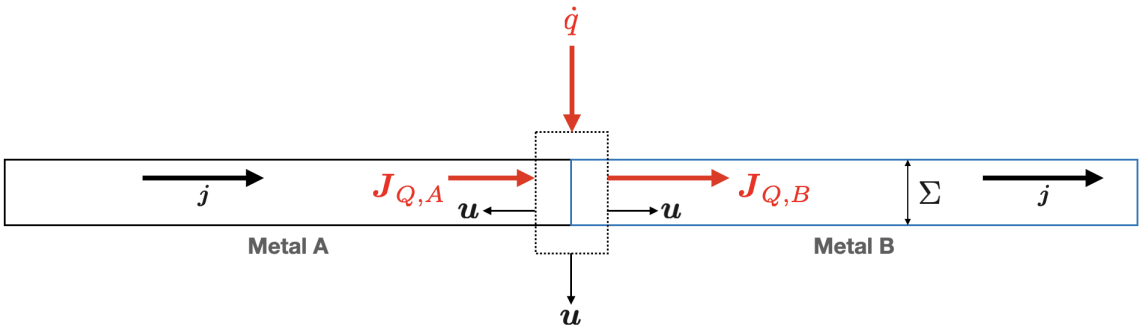



FIGURE 12.2 – Peltier effect

In a stationary regime, the electric current is uniform, and is therefore the same in both regions. At the junction, the heat fluxes associated to the Peltier effect are different on each side

$$\vec{J}_{Q,A} = \Pi_A \vec{j} \qquad \vec{J}_{Q,B} = \Pi_B \vec{j} \tag{12.52}$$

By considering a system consisting of the metal junction, delimited by a cylindrical imaginary surface, we can write the amount of heat gained by the system, through exchanges along the wire and with the thermostat through the lateral surfaces. We write $\dot{q}$ for the heat gained by the wire from the environment, through these lateral surfaces, per unit time. We notice that the heat gained per unit time on the left side is $\Pi_A j \Sigma$, where $\Sigma$ is the section of the wire, and the heat gained per unit time on the right side is $-\Pi_B j \Sigma$. In a stationary regime, the total heat gained must be null, so

$$\dot{q} + (\Pi_A - \Pi_B)\, j\Sigma = 0 \tag{12.53}$$

The heat output, gained by the environment at the junction, is then

$$\dot{Q} = -\dot{q} = (\Pi_A - \Pi_B)\, j\Sigma \tag{12.54}$$

Depending on the thermopowers of $A$ and $B$, this may be an actual heating of the environment $\dot{Q} > 0$, or it may be cooling $\dot{Q} < 0$. This is actually what is put to use for localized cooling using so-called Peltier modules on CPUs and GPUs. As we shall see below, since it must compete with Joule heating that goes as $j^2$, some optimal electric current $j$ is required for best cooling.

### Joule effect

Let us consider a homogenenous conducting wire at uniform temperature through which an electric current passes, in a stationary regime. The heat flux in this case is, as we saw, $\vec{J}_Q = -e\varepsilon T \vec{J}_n$ so the energy density current is

$$\vec{J}_e = \vec{J}_Q + \mu \vec{J}_n = (\mu - e\varepsilon T)\, \vec{J}_n \tag{12.55}$$

where the chemical potential is given by $\mu = \mu_0 - eV$. The divergence of the energy density current, appearing in the energy conservation equation, is then

$$\vec{\nabla} \cdot \vec{J}_e = (\mu - e\varepsilon T)\, \vec{\nabla} \cdot \vec{J}_n + \vec{J}_n \cdot \vec{\nabla}\, (\mu - e\varepsilon T) \tag{12.56}$$

The first term vanishes due to the conservation of particles and the stationary hypothesis, because $\vec{\nabla} \cdot \vec{J}_n = 0$. With the uniform temperature assumption and the homogeneity of the wire, the second term also simplifies and we have

$$\vec{\nabla} \cdot \vec{J}_e = \vec{J}_n \cdot \vec{\nabla}\mu = -e\vec{J}_n \cdot \vec{\nabla}V = -\vec{j} \cdot \vec{\mathcal{E}} \tag{12.57}$$

Considering a small volume of the wire, it exchanges energy with the rest of the wire and with the thermostat. The local energy conservation then reads

$$\frac{\partial e}{\partial t} = -\vec{\nabla} \cdot \vec{J}_e + \dot{q} = 0 \tag{12.58}$$

where the first term in the middle expression is the energy gained by the wire volume through the sections on the left and right, and the second term is the heat gained from the thermostat maintaining the temperature constant, through the lateral surfaces of the wire. The nullity of the sum of these two terms is because of the stationarity hypothesis. We therefore have

$$\dot{q} = -\vec{j} \cdot \vec{\mathcal{E}} = -\sigma \mathcal{E}^2 = -\frac{j^2}{\sigma} \tag{12.59}$$

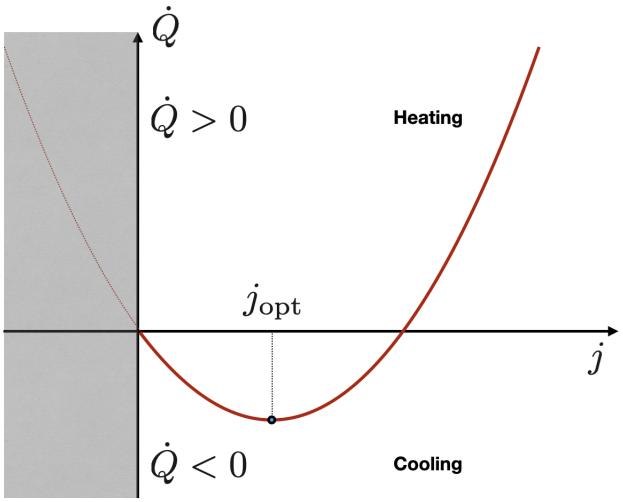



FIGURE 12.3 – Competition between the Joule and Peltier effects, showing the optimal electric current.

This quantity $\dot{q} < 0$ is the heat gained by the wire volume element from the environment, so the heat is actually gained by this environment, i.e. the thermostat, since it is $-\dot{q} > 0$. This is the **Joule effect**. At a junction, this effect occurs simultaneously with the Peltier effect, so the heat output, gained by the environment, is

$$\dot{Q} = \frac{j^2}{\sigma} + (\Pi_A - \Pi_B)\, j\Sigma \tag{12.60}$$

If $\Pi_A < \Pi_B$ we have actual cooling of the environment at low current $\dot{Q} < 0$, but it becomes net heating at larger currents, $\dot{Q} > 0$, as shown in Fig. 12.3. Note also that in the presence of a temperature gradient, a supplementary, weak **Thomson effect** occurs.

# 13

# Evolution towards equilibrium in classical thermodynamics

As we pointed out in the beginning, thermodynamics of systems at equilibrium describes macroscopic systems through a small number of macroscopic variables, and the trade-off for this is that it is necessary to introduce the entropy, a quantity that has no microscopic equivalent but describes our ignorance of the microscopic details of the system. In this chapter, we discuss the thermodynamic functions and potentials, the conditions for equilibrium and the evolution of a system in a fixed environment. Some of the results of this chapter have already been established before, but they are shown here in a new light.

## 13.1 Thermodynamic functions

### 13.1.1 The fundamental postulate and entropy

From primary state variables describing an isolated system (energy $E$, volume $V$, number of particles $N$, and any other variable $x$), we postulate the existence of a function $S(E, V, N, x)$, the entropy, such that its knowledge suffices to know all the (macroscopic) properties of the system at equilibrium. The properties of this function are :

— It must be continuous and derivable
— Its derivative with respect to energy is positive $\partial S/\partial E > 0$
— When a constraint is relaxed in an isolated system, the entropy can only increase ($S_f \geqslant S_i$)
— For weakly coupled sub-systems, it is additive ($S = S_1 + S_2$). This supposes that no long-range interaction exists, however. For electromagnetic forces, screening in globally neutral systems effectively prevents this from happening, but it is not the case in gravitational systems.

The derivation of $S$ with respect to its different variables defines the temperature $T$, pressure $P$, chemical potential $\mu$, and the thermodynamic conjugate $X$ of $x$ through

$$\frac{\partial S}{\partial E} = \frac{1}{T} \qquad \frac{\partial S}{\partial V} = \frac{P}{T} \qquad \frac{\partial S}{\partial N} = -\frac{\mu}{T} \qquad \frac{\partial S}{\partial x} = \frac{X}{T} \tag{13.1}$$

The differential of $S$ therefore reads, as is now well known,

$$\mathrm{d}S = \frac{1}{T}\mathrm{d}E + \frac{P}{T}\mathrm{d}V - \frac{\mu}{T}\mathrm{d}N + \frac{X}{T}\mathrm{d}x \tag{13.2}$$

The thermodynamic conjugates $T$, $P$, $\mu$, and $X$ are functions of the same variables as $S$, that is $E$, $V$, $N$, and $x$, and the **Maxwell relations** expressing the identity of cross double derivatives hold, e.g.,

$$\frac{\partial}{\partial V}\left(\frac{1}{T}\right) = \frac{\partial^2 S}{\partial V \partial E} = \frac{\partial^2 S}{\partial E \partial V} = \frac{\partial}{\partial E}\left(\frac{P}{T}\right) \tag{13.3}$$

### 13.1.2 Internal energy

Let us assume that the system is globally at rest, so that its energy $E$ is just the internal energy $U$. Since $\partial S/\partial E = \partial S/\partial U > 0$, the relation between the entropy $S$ and the (internal) energy $U$ is one-on-one. The relation $S(U, V, N, x)$ may then be (at least formally) inverted into $U(S, V, N, x)$ and the differential of the energy reads

$$\mathrm{d}U = T\mathrm{d}S - P\mathrm{d}V + \mu\mathrm{d}N - X\mathrm{d}x \tag{13.4}$$

where $T$, $P$, $\mu$, and $X$ may be viewed as functions of the "new" variables $S$, $V$, $N$, and $x$. In this expression, the quantity $T\mathrm{d}S$ is the heat exchange and the sum $-P\mathrm{d}V - X\mathrm{d}x$ is the reversible work exchange. From the above relation we may also write Maxwell relations such as

$$\left(\frac{\partial T}{\partial V}\right)_{S,N,x} = -\left(\frac{\partial P}{\partial S}\right)_{V,N,x} \tag{13.5}$$

Both expressions of $\mathrm{d}U$ and $\mathrm{d}S$ are valid in any case, and the two representations are equivalent. What differs is that $S$ is the "right" thermodynamic function to use when the variables to consider are $(U, V, N, x)$, while $U$ is the "right" thermodynamic function to use when the variables to consider are $(S, V, N, x)$.

### 13.1.3 Legendre transformations

Now what is the "right" thermodynamic function to use when the variables to consider are different from these two examples, say for instance $(T, V, N, x)$ ? It must be one whose differential is easily expressed in terms of the increments of these variables. Such is the **Helmholtz free-energy** that we already encountered, and that is defined through a **Legendre transformation** from the internal energy

$$F = U - TS \tag{13.6}$$

Its differential is then clearly expressed in terms of the increments of $(T, V, N, x)$

$$\mathrm{d}F = \mathrm{d}U - T\mathrm{d}S - S\mathrm{d}T = -S\mathrm{d}T - P\mathrm{d}V + \mu\mathrm{d}N - X\mathrm{d}x \tag{13.7}$$

where $S$, $P$, $\mu$, and $X$ are now functions of $(T, V, N, x)$ and various Maxwell relations can be derived between them. It is left to the reader to derive some of those.

It should be noted that $F(T, V, N, x)$ holds all the information needed to describe the system. Indeed, from it we get the entropy $S(T, V, N, x)$ via

$$S(T, V, N, x) = -\left(\frac{\partial F}{\partial T}\right)_{V,N,x} \tag{13.8}$$

and the internal energy from the inverse Legendre transformation

$$U(T, V, N, x) = F(T, V, N, x) + TS(T, V, N, x) \tag{13.9}$$

Finally, from $S(T, V, N, x)$ and $U(T, V, N, x)$ we may eliminate $T$ and get, at least formally, $S(U, V, N, x)$. The fundamental postulate then ensures the result. More generally, any Legendre transformation such as the one above will exchange a variable with its conjugate, yielding a new function that also holds

all the information. We thus define the **enthalpy** $H(S,P,N,x)$, the **Gibbs free-energy** $G(T,P,N,x)$, and the **grand-canonical potential** $J(T,V,\mu,x)$ through the following relations

$$H = U + PV \qquad G = U - TS + PV \qquad J = U - TS - \mu N \tag{13.10}$$

### 13.1.4 Example of the simple fluid

A simple fluid is described by three state variables[1], and if we fix the number of particles $N$, its state is now described by only two independent variables (one mechanical, and one thermodynamical). Depending on the choice of these variables, the "right" function to use is different.

| Functions | Variables | Conjugates | Maxwell relations |
|---|---|---|---|
| $U$ | $(S,V)$ | $T = \frac{\partial U}{\partial S} \quad P = -\frac{\partial U}{\partial V}$ | $\left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial P}{\partial S}\right)_V$ |
| $H$ | $(S,P)$ | $T = \frac{\partial H}{\partial S} \quad V = \frac{\partial H}{\partial P}$ | $\left(\frac{\partial T}{\partial P}\right)_S = \left(\frac{\partial V}{\partial S}\right)_P$ |
| $F$ | $(T,V)$ | $S = -\frac{\partial F}{\partial T} \quad P = -\frac{\partial F}{\partial V}$ | $\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial P}{\partial T}\right)_V$ |
| $G$ | $(T,P)$ | $S = -\frac{\partial G}{\partial T} \quad V = \frac{\partial G}{\partial P}$ | $\left(\frac{\partial S}{\partial P}\right)_T = -\left(\frac{\partial V}{\partial T}\right)_P$ |

### 13.1.5 Example use of Maxwell relations

As an example of the use of Maxwell relations, we consider the computation of the force between the two plates of a capacitor. We write $\vec{f}$ for the force exerted by an operator to maintain the system at equilibrium (i.e. to compensate the attraction between the two oppositely-charged plates). The system under consideration is the capacitor, and the variation of its internal energy is

$$\mathrm{d}U = T\mathrm{d}S + \Phi \mathrm{d}q + f\mathrm{d}x \tag{13.11}$$

where $\Phi$ is the electrostatic potential between the plates, $q$ the electric charge on one of the plates, and $x$ the spacing between the two plates. The last two terms on the right-hand side above represent the reversible work of external forces.

Let us first choose $(T,x,q)$ as variables. The proper thermodynamic function to use is the Helmholtz free-energy $F = U - TS$ such that $\mathrm{d}F = -S\mathrm{d}T + \Phi\mathrm{d}q + f\mathrm{d}x$. Maxwell relations derived from this expression include, in particular,

$$\left(\frac{\partial f}{\partial q}\right)_{T,x} = \left(\frac{\partial \Phi}{\partial x}\right)_{T,q} \tag{13.12}$$

Now the electrostatic potential is, assuming no edge-effects (i.e., that the plates are effectively inifinite),

$$\Phi = \frac{q}{C} = \frac{xq}{\epsilon_0 \epsilon_r(T)\Sigma} \tag{13.13}$$

with $\Sigma \gg x^2$ the surface of each plate. We then have the derivative of the force

$$\left(\frac{\partial f}{\partial q}\right)_{T,x} = \frac{q}{\epsilon_0 \epsilon_r(T)\Sigma} \tag{13.14}$$

1. This means there is no $x$ variable as used in the previous paragraphs.

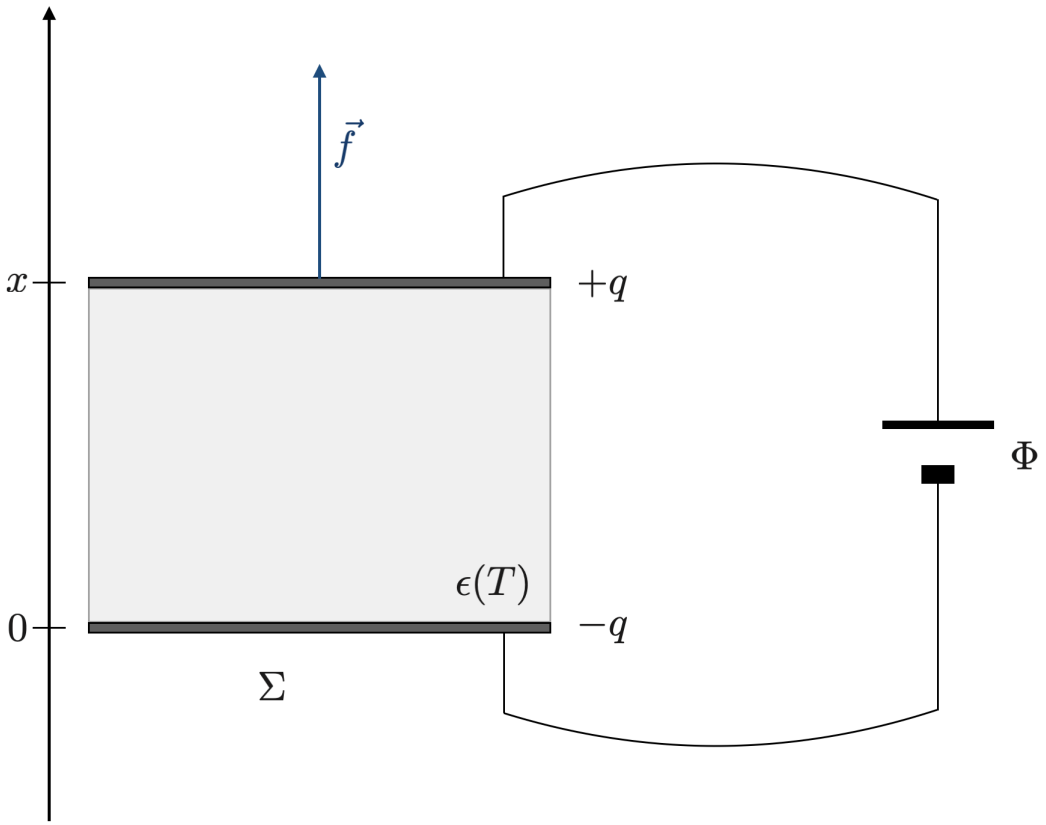



FIGURE 13.1 – A capacitor

and from there the force itself, as a funtion of the requested variables

$$f(T, x, q) = \frac{q^2}{2\epsilon_0 \epsilon_r(T)\Sigma} \tag{13.15}$$

The "constant" that should appear is in fact null because the force must vanish when the charge goes to zero. We leave it to the reader to show that, had we chosen another set of variables, e.g. $(T, x, \Phi)$, we would have found the equivalent result

$$f(T, x, \Phi) = \frac{\epsilon_0 \epsilon_r(T)\Sigma\Phi^2}{2x^2} \tag{13.16}$$

### 13.1.6 Equation of state and Maxwell relations

The equation of state relates the various macroscopic state variables at equilibrium. For a simple fluid, this takes the form $f(P, V, T, N) = 0$. From this equation, and using Maxwell relations, it is possible to access some of the derivatives of the entropy. For instance, using $F$ we have

$$\left(\frac{\partial S}{\partial V}\right)_{T,N} = \left(\frac{\partial P}{\partial T}\right)_{V,N} \tag{13.17}$$

and the right-hand side may be determined directly from the equation of state. Similarly, from $G$ we have

$$\left(\frac{\partial S}{\partial P}\right)_{T,N} = -\left(\frac{\partial V}{\partial T}\right)_{P,N} \tag{13.18}$$

whose right-hand side may also be determined directly from the equation of state. More generally, the dependence of $S$ on variables other than $T$ may be thus determined, but the way $S$ depends on $T$ cannot be computed this way. These must be determined from statistical physics, as we did earlier. Similar considerations apply to the determination of the properties of the heat capacities

$$C_v = \left(\frac{\partial U}{\partial T}\right)_V = T\left(\frac{\partial S}{\partial T}\right)_V \qquad C_p = \left(\frac{\partial H}{\partial T}\right)_P = T\left(\frac{\partial S}{\partial T}\right)_P \tag{13.19}$$

The dependence of these on the mechanical variables are determined from the equation of state

$$\left(\frac{\partial C_v}{\partial V}\right)_T = T\left(\frac{\partial^2 P}{\partial^2 T}\right)_V \qquad \left(\frac{\partial C_p}{\partial P}\right)_T = -T\left(\frac{\partial^2 V}{\partial^2 T}\right)_P \tag{13.20}$$

but the partial derivatives of $C_v$ and $C_p$ with respect to temperature cannot be determined from the equation of state.

### 13.1.7 Extensivity and intensivity

By definition $V$ and $N$ are extensive variables. If weak coupling conditions apply, so are the energy $U$ and the entropy $S$. On the other hand, $T$, $P$, and $\mu$ are intensive. For a simple fluid, the extensivity of $U$ and $S$ may be written mathematically as $U(\lambda S, \lambda V, \lambda N) = \lambda U(S, V, N)$ for any $\lambda$. This is stated as $U$ being a **first-degree homogeneous function** of its variables. By deriving this relation with respect to $\lambda$ and expressing the result for $\lambda = 1$, we obtain the **Euler relation**

$$S\left(\frac{\partial U}{\partial S}\right)_{V,N} + V\left(\frac{\partial U}{\partial V}\right)_{S,N} + N\left(\frac{\partial U}{\partial N}\right)_{S,V} = U \tag{13.21}$$

This yields the following simple expression for the internal energy, and from there, for the various thermodynamic functions using their definitions as Legendre transforms

$$U = TS - PV + \mu N \tag{13.22}$$

$$H = TS + \mu N \qquad F = -PV + \mu N \qquad J = -PV \qquad G = \mu N \tag{13.23}$$

From this we may derive the **Gibbs-Duhem relation** for a pure compound (i.e., not a mixture), which we leave for the reader to derive (starting from the differential of the Gibbs free-energy),

$$-S\mathrm{d}T + V\mathrm{d}P = N\mathrm{d}\mu \tag{13.24}$$

The chemical potential $\mu$ is therefore a function of $(T, P)$ and we have

$$\left(\frac{\partial \mu}{\partial T}\right)_P = -\frac{S}{N} = -s \qquad \left(\frac{\partial \mu}{\partial P}\right)_T = \frac{V}{N} = v \tag{13.25}$$

More generally, an intensive quantity can only depend on intensive variables, as is the case here for $\mu(T, P)$. For instance, $P(S, V, N)$ being intensive can in fact only depend on $s = S/N$ and $v = V/N$.

## 13.2 Nernst's principle

### 13.2.1 Necessity of the third principle of thermodynamics

In a classical approach, the energy is defined up to an additive constant $U_0$ (which must be proportional to $N$ to preserve extensivity), so the chemical potential is also defined up to an additive

constant. Since experiments only give access to energy variations $\Delta U$, this is not a problem. Entropy is also *a priori* defined up to an additive constant $S_0$, since what is measured in a given experiment is the variation $\Delta S$. However, this poses a problem for the Helmholtz and Gibbs free energies, $F = U - TS$ and $G = U - TS + PV$ since these are now defined up to the addition of $U_0 - TS_0$ and this is not a constant if the temperature varies between the initial and final states of the system. In that case, the variations $\Delta F$ and $\Delta G$ are useless.

### 13.2.2 The third principle of thermodynamics

As we saw in earlier chapters, a statistical physics approach allows to determine $S$ from first principles (Boltzmann's formula $S = k_B \ln \Omega$ taking into account all degrees of freedom). In classical thermodynamics, this is not possible and is supplemented by **Nernst's principle**, also known as the **third principle**, which states that for fixed $(V, N)$ or $(P, N)$,

$$\lim_{T\to 0} S(T, V, N) = 0 \qquad \lim_{T\to 0} S(T, P, N) = 0 \tag{13.26}$$

From a statistical point of view, this would be stating that the fundamental state is non-degenerate at $T = 0$, $\Omega_0 = 1$. In practice, $\Omega_0$ is small, so that $S(T = 0) \approx 0$ for all intents and purposes.

### 13.2.3 Consequences

As a consequence, $S$ is no longer defined up to an additive constant, i.e. $S_0 = 0$, which solves the problem of the indetermination of $\Delta F$ and $\Delta G$. Moreover, since $S(T, V) \to S(0, V) = S_0 = 0$ and $S(T, P) \to S(0, P) = S_0 = 0$ become independent of $V$ and $P$ (respectively) in the low-temperature limit, we have, from Maxwell's relations

$$\left(\frac{\partial P}{\partial T}\right)_V = \left(\frac{\partial S}{\partial V}\right)_T \to 0 \qquad \left(\frac{\partial V}{\partial T}\right)_P = -\left(\frac{\partial S}{\partial P}\right)_T \to 0 \tag{13.27}$$

in the limit $T \to 0$. These forms of the entropy may be computed from the heat capacities

$$S(T, V) = \int_0^T \frac{C_v(T', V)}{T'} \mathrm{d}T' \qquad S(T, P) = \int_0^T \frac{C_p(T', P)}{T'} \mathrm{d}T' \tag{13.28}$$

and since these must remain finite as $T \to 0$, both heat capacities must also go to zero [2],

$$\lim_{T\to 0} C_v(T, V) = 0 \qquad \lim_{T\to 0} C_p(T, P) = 0 \tag{13.29}$$

Another consequence is that it is impossible to reach $T = 0$ through an adiabatic process. Indeed, if the process is irreversible, $\Delta S = S_f - S_i > 0$ but the final entropy is null $S_f = 0$ since the final state has $T = T_f = 0$ and the initial entropy cannot be negative. If the process is reversible, $\Delta S = 0$ but the conclusion stands. Should we consider a process that is not adiabatic, it is less direct to show but the result stands, leading to the impossibility to reach absolute zero.

2. for example, take an ansatz that $C_v(T) \propto T^\alpha$ with $\alpha \neq 0$. Then we find that $S \propto T^\alpha \propto C_v$, which proves the result that $C_v$ must go to zero as $T \to 0$.

## 13.3 Physical interpretation of thermodynamic potentials

### 13.3.1 Internal energy

Let us consider a system with a fixed number of particles $N$. The first principle of thermodynamics applied to this system in a given process states that $\Delta U = Q + W$. In particular, in an adiabatic process for an isolated system, we have $\Delta U_{\rm ad} = W_{\rm ad}$. The external medium receives an amount of work $\mathcal{T} = -W$. Consequently, the amount of work received by the external medium in an adiabatic process (whether reversible or not), equals the decrease of the system's internal energy

$$\mathcal{T}_{\rm ad} = -W_{\rm ad} = -\Delta U_{\rm ad} \tag{13.30}$$

### 13.3.2 Enthalpy

Let us now consider a system that is in contact with a volume reservoir, imposing a given constant pressure $P_0$. We define the quantity

$$H_0 = U + P_0 V \tag{13.31}$$

It should be stressed that $H_0$ is not the enthalpy, which is defined by $H = U + PV$ where the pressure is that of the system, not of the reservoir. For any process, this quantity varies by an amount

$$\Delta H_0 = \Delta U + P_0 \Delta V = Q + W + P_0 \Delta V \tag{13.32}$$

but the amount of work received by the system includes both that related to the variation of the volume[3] $-P_0\Delta V$ and any other work $W'$, so $W = -P_0\Delta V + W'$. Finally, we have

$$\Delta H_0 = Q + W' \tag{13.33}$$

In an adiabatic process, we have $\Delta H_{0,\rm ad} = W'_{\rm ad}$ and thus the amount of work (other than that associated to volume variations) received by the external medium in an adiabatic process at constant external pressure is equal to the decrease of $H_0$ in that process

$$\mathcal{T}_{\rm ad} = -\Delta H_{0,\rm ad} \tag{13.34}$$

It is only in the case that the initial and final pressure of the system are equal to $P_0$ that $\Delta H = \Delta H_0$. A counter-example is that shown in Fig. 13.2. If we release the lock, the final pressure is $P_f = P_0$, while the initial pressure was much larger. In that case $\Delta H \neq \Delta H_0$.

### 13.3.3 Helmholtz free energy

Consider now a system that is only allowed to exchange heat with a thermostat imposing its temperature $T_0$. Much like before, we define the quantity

$$F_0 = U - T_0 S \tag{13.35}$$

which is not equal to the Helmholtz free energy $F = U - TS$. In any process, the variation of this quantity is

$$\Delta F_0 = \Delta U - T_0 \Delta S = W + Q - T_0 \Delta S \tag{13.36}$$

3. It may be useful to recall that the infinitesimal work received by the system due to pressure forces is $\delta W = -P_0 {\rm d}V$ since it is the external medium that does this pressure work.

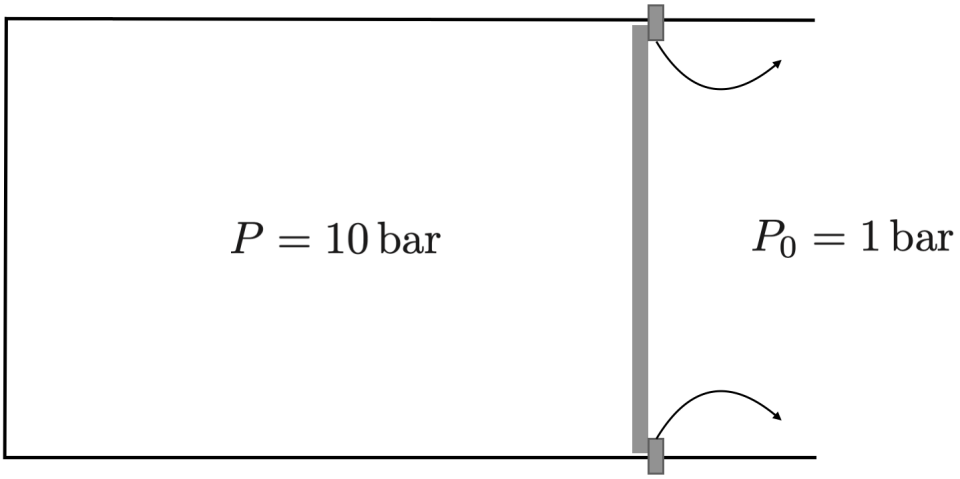



FIGURE 13.2 – Releasing an overpressured gas

and the second principle states that $\Delta S \geqslant Q/T_0$ with equality if and only if the process is reversible. Consequently, $Q \leqslant T_0 \Delta S$ and

$$\Delta F_0 \leqslant W \tag{13.37}$$

The amount of work received by the external medium is then at most equal to the decrease of $F_0$, and it reaches this maximum when the transformation is reversible

$$\mathcal{T} = -W \leqslant -\Delta F_0 \tag{13.38}$$

This explains the name of "free energy" : it is the maximum amount of work that can be extracted from the system. We note that in the case $T_f = T_i = T_0$, the variation of $F_0$ equals that of $F$, and then $\mathcal{T} \leqslant -\Delta F$.

### 13.3.4 Gibbs free energy

Consider now a system exchanging heat with a thermostat at temperature $T_0$ and volume with a reservoir at pressure $P_0$. We define, similarly to the two previous cases, the function

$$G_0 = U - T_0 S + P_0 V \tag{13.39}$$

which is not to be confused with the system's Gibbs free energy $G = U - TS + PV$. The variation of $G_0$ in any process reads

$$\Delta G_0 = \Delta U - T_0 \Delta S + P_0 \Delta V = Q + W - T_0 \Delta S + P_0 \Delta V = Q + W' - T_0 \Delta S \tag{13.40}$$

Using the second principle as we did above, we get that the amount of work (other than that associated to volume variations) received by the external medium is then at most equal to the decrease of $G_0$, and it reaches this maximum when the transformation is reversible

$$\mathcal{T}' = -W' \leqslant -\Delta G_0 \tag{13.41}$$

In the case $T_f = T_i = T_0$ and $P_f = P_i = P_0$, we have $\Delta G = \Delta G_0$ and therefore $\mathcal{T}' \leqslant -\Delta G$,

with the equality reached for a reversible process, in which case the amount of (non-volume related) work extracted from the system is maximum, $\mathcal{T}' = -\Delta G$. This will in particular apply to any system that is placed in contact with the atmosphere (which imposes its pressure and temperature), and may be used to derive the amount of work extracted that is not related to volume variations, such as electric work in a battery or the discharge of a capacitor.

## 13.4 Evolution and equilibrium conditions

### 13.4.1 General conditions

We consider the evolution from an initial state $i$ to a final state $f$, that are both equilibria. The two principles of thermodynamics state that

$$\Delta U = Q + W \qquad\qquad \Delta S \geqslant \int_i^f \frac{\delta Q}{T_{\text{ext}}} \tag{13.42}$$

These two conditions must be met by any physically sound evolutionary path. In the specific case of an isolated system, we have $\Delta U = 0$ and $\Delta S \geqslant 0$, so $S_f \geqslant S_i$ and the final state corresponds to the maximum value of the entropy given the remaining constraints. This is a result we saw already.

### 13.4.2 Equilibrium between two sub-systems of an isolated system

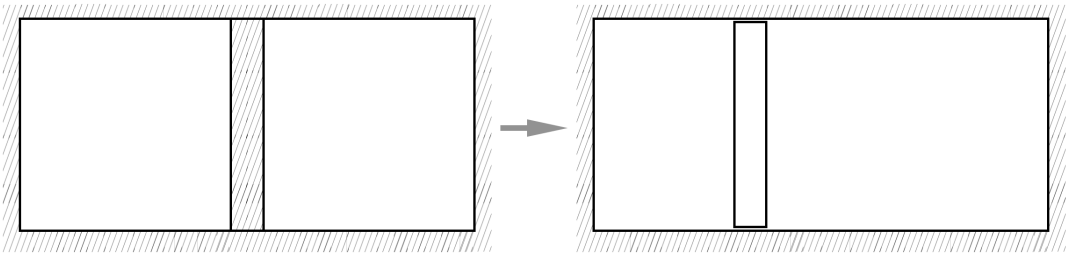


FIGURE 13.3 – Releasing constraints, from an initial state (left) to a final one (right)

Let us consider an isolated system consisting of two sub-systems $A$ and $B$, that are initially isolated from each other. Their internal energies, volumes, and numbers of particles are respectively $(U_A, V_A, N_A)$ and $(U_B, V_B, N_B)$. At $t = 0$, we imagine that the panel separating the two systems is rendered mobile and allowed to let heat through, and we wait for a new, final equilibrium to settle. The external parameters are the total energy $U_0 = U_A + U_B$, the total volume $V_0 = V_A + V_B$, and the numbers of particles in each sub-system $N_A$ and $N_B$. The energies and volumes of each sub-system are now internal variables, and their values are determined in the final equilibrium in such a way that the entropy of $A \cup B$ is maximum. This entropy is given by

$$S(U_A, V_A; U_0, V_0, N_A, N_B) = S_A(U_A, V_A) + S_B(U_0 - U_A, V_0 - V_A) \tag{13.43}$$

and must reach a maximum with respect to the two independent variables $U_A$ and $V_A$, so

$$\frac{\partial S}{\partial U_A} = \frac{\partial S_A}{\partial U_A} - \frac{\partial S_B}{\partial U_B} = \frac{1}{T_A} - \frac{1}{T_B} = 0 \tag{13.44}$$

$$\frac{\partial S}{\partial V_A} = \frac{\partial S_A}{\partial V_A} - \frac{\partial S_B}{\partial V_B} = \frac{P_A}{T_A} - \frac{P_B}{T_B} = 0 \tag{13.45}$$

The equilibrium is then attained when the subsystems share the same temperature and pressure,

$$T_A(U_A, V_A, N_A) = T_B(U_0 - U_A, V_0 - V_A, N_B) \tag{13.46}$$

$$P_A(U_A, V_A, N_A) = P_B(U_0 - U_A, V_0 - V_A, N_B) \tag{13.47}$$

These are two equations for the two unknowns $U_A$ and $V_A$, so these can be solved, at least formally, and the entropy at equilibrium obtained by reinjecting these into the expression for $S$ above. Note that if the panel is only allowed to let heat through but not to move, there is only one equation, $T_A = T_B$, but also only one unknown ($U_A$). The panel can also be made porous, in which case $N_A$ and $N_B$ become internal variables, and the equilibrium is found through

$$\frac{\partial S}{\partial N_A} = \frac{\partial S_A}{\partial N_A} - \frac{\partial S_B}{\partial N_B} = -\frac{\mu_A}{T_A} + \frac{\mu_B}{T_B} = 0 \tag{13.48}$$

expressing the equality of chemical potentials. Finally, it should also be noted that the maximum entropy may not correspond to $\mathrm{d}S = 0$, as is the case for the Joule and Gay-Lussac experiment shown in Fig. 13.4. There, the external parameters are $U_0$ and $V_0$ and the internal variable is the volume $V$ occupied by the gas. Since

$$\left(\frac{\partial S}{\partial V}\right)_U = \frac{P}{T} > 0 \tag{13.49}$$

the maximum entropy is reached for the maximum volume and the final volume is $V_f = V_0$.

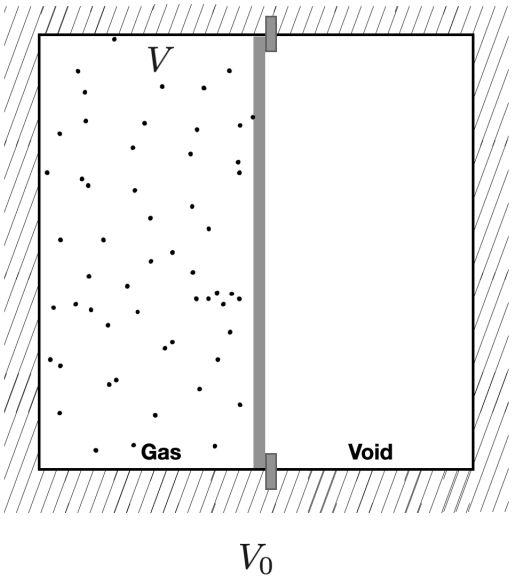



FIGURE 13.4 – The Joule and Gay-Lussac experiment

To summarize globally, for an isolated system, external parameters such as $U$, $V$, $x$, ... are fixed and if we consider an internal variable $y$, the equilibrium is reached when $S(y; U, V, x, \ldots)$ is maximum,

$$\left(\frac{\partial S}{\partial y}\right)_{U,V,x,\ldots} = 0 \tag{13.50}$$

yielding the equilibrium value $y_{\mathrm{eq}}(U, V, x, \ldots)$ of the internal variable $y$. Note that for this to be a maximum we should have

$$\left(\frac{\partial^2 S}{\partial y^2}\right)_{U,V,x,\ldots} < 0 \tag{13.51}$$

### 13.4.3 Maximum entropy and minimum energy

The maximum of the entropy with respect to $y$ for constant $U$, $V$, $N$, $x$, ... is equivalent to $U(y; S, V, N, x, \ldots)$ reaching a minimum with respect to $y$ for constant $S$, $V$, $N$, $x$, .... Mathematically,

$$\left(\frac{\partial S}{\partial y}\right)_{U,V,N,x,\ldots} = 0 \qquad \left(\frac{\partial^2 S}{\partial y^2}\right)_{U,V,N,x,\ldots} < 0 \tag{13.52}$$

is equivalent to the alternative formulation of equilibrium

$$\left(\frac{\partial U}{\partial y}\right)_{S,V,N,x,\ldots} = 0 \qquad \left(\frac{\partial^2 U}{\partial y^2}\right)_{S,V,N,x,\ldots} > 0 \tag{13.53}$$

This is best understood graphically, as shown in Fig. 13.5, taken from Callen's "Thermodynamics". Indeed, let us consider a fixed internal energy $U$. Equilibrium occurs for the maximum value of $S$ as a function of other variables (here just one, that we call $x$), so the section of the surface $S(U, x)$ at constant $U$, is a concave function of $x$. Now, since $\partial S/\partial U = 1/T > 0$, these functions of $x$ must increase for increasing values of $U$. The surface $S(U, x)$ then necessarily has the shape shown in Fig. 13.5. By performing a cut at constant $S$, one sees that the equilibrium point corresponds to the minimum $U$.

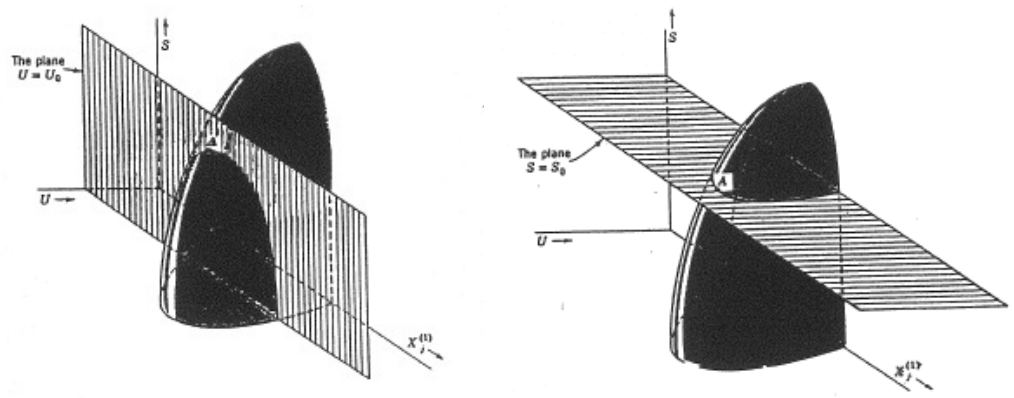



FIGURE 13.5 – Equivalence between the equilibrium conditions that $S$ should be maximum at fixed $U$, and that $U$ should be minimum at fixed $S$.

### 13.4.4 Evolution and equilibrium in a fixed environment

**Minimum of $G_0$**

We consider a system $\mathcal{S}$ exchanging heat $Q$ with a thermostat $\mathcal{T}$ imposing its temperature $T_0$, and work $W$ with a reservoir of volume $\mathcal{R}$ imposing pressure $P_0$. The variation of the system's entropy in any process is such that

$$\Delta S = S_f - S_i \geqslant \frac{Q}{T_0} \tag{13.54}$$

and the work received by the system is $W = -P_0 \Delta V$. The variation of its internal energy is then

$$\Delta U = Q + W \leqslant T_0 \Delta S - P_0 \Delta V \tag{13.55}$$

so that if we define $G_0 = U - T_0 S + P_0 V$ we have

$$\Delta G_0 \leqslant 0 \tag{13.56}$$

Any physically allowed process is such that it can only lead to a decrease of $G_0$, and the final equilibrium corresponds to a state of minimal $G_0$ compatible with any remaining constraint on the system.

With fixed external parameters $P_0$, $T_0$, and $N$, and internal variables "chosen" by the system $\mathcal{S}$ (either through exchanges with the environment, such as $U$ and $V$ or through internal exchanges, $y$), the differential of $G_0$ reads

$$\mathrm{d}G_0 = \mathrm{d}U - T_0 \mathrm{d}S + P_0 \mathrm{d}V = \mathrm{d}U - T_0 \left( \frac{1}{T} \mathrm{d}U + \frac{P}{T} \mathrm{d}V + \frac{Y}{T} \mathrm{d}y \right) + P_0 \mathrm{d}V \tag{13.57}$$

where we have inserted the expression of $\mathrm{d}S$ based on its natural variables. Rearranging, we get

$$\mathrm{d}G_0 = \left( 1 - \frac{T_0}{T} \right) \mathrm{d}U + \left( P_0 - \frac{T_0}{T} P \right) \mathrm{d}V - \frac{T_0}{T} Y \mathrm{d}y \tag{13.58}$$

In order for $G_0$ to be minimum as a function of the internal variables of the system, we must have

$$T = T_0 \qquad P = P_0 \qquad Y = 0 \tag{13.59}$$

which constitute three equations for three unknowns ($U$, $V$, and $y$). In practice, we may use $S$, $V$, and $y$ as internal variables, so we would replace $\mathrm{d}U$ in the first expression of $\mathrm{d}G_0$, yielding

$$\mathrm{d}G_0 = (T - T_0)\, \mathrm{d}S + (P_0 - P)\, \mathrm{d}V - Y \mathrm{d}y \tag{13.60}$$

and therefore leading to the same three conditions for equilibrium.

### Specific cases

We may consider a few specific cases. For instance, in an isolated rigid system, $U$ and $V$ are fixed, so $G_0 = U - T_0 S + P_0 V$ minimum corresponds to $-T_0 S$ minimum, i.e., $S$ maximum. In a system that is only in thermal contact with $\mathcal{T}$ but has a fixed volume, the condition for equilibrium is equivalent to $F_0 = U - T_0 S$ being minimum. This leads to the equation $T = T_0$ and so at equilibrium $F_0$ identifies with $F = U - TS$. Its dependence on the internal variable $y$ may be explicitly stated as $F(y; T, V)$, and the determination of the value of the internal variable $y$ at equilibrium may then be done through

$$Y = \left( \frac{\partial F}{\partial y} \right)_{T,V} = 0 \tag{13.61}$$

leading to $y_{\mathrm{eq}}(T, V)$ and the Helmholtz free-energy at equilibrium is

$$F(T, V) = F\left[ y_{\mathrm{eq}}(T, V); T, V \right] \tag{13.62}$$

The same reasoning applies to a system in thermal contact with $\mathcal{T}$ ($T_0$) and mechanical contact with $\mathcal{R}$ ($P_0$). In that case, at equilibrium $T = T_0$ and $P = P_0$ so that $G_0$ identifies with $G = U - TS + PV$. Its dependence on the internal variable $y$ may be explicitly stated as $G(y; T, P)$ and the value of this variable at equilibrium found through

$$Y = \left( \frac{\partial G}{\partial y} \right)_{T,P} = 0 \tag{13.63}$$

leading to $y_{\mathrm{eq}}(T, P)$ and the Gibbs free-energy at equilibrium is

$$G(T, P) = G\left[ y_{\mathrm{eq}}(T, P); T, P \right] \tag{13.64}$$

**Remarks**

We note that to find the conditions for equilibrium, the choice of external parameters is a matter of taste, because at equilibrium, everything is fixed. This is akin to the thermodynamic limit discussed earlier. Once this choice is made, the function to be minimized with respect to the internal variable $y$ is unique.

Now if we study the irreversible evolution after relaxing a constraint, the choice of external parameters is not free, but dictated by the experimental conditions. For instance, in the Joule and Gay-Lussac experiment, if the system is isolated we should work with $S$, as we know $S_f > S_i$, but if the system is in contact with a thermostat, we should work with $F$ as $F_f < F_i$.

Finally, if there are other exchanges possible (for instance particles with a reservoir imposing its chemical potential $\mu_0$, or work in a gravitational field, or charges with a generator of fixed electromotive force, ...) all this reasoning should be redone from the start.

### 13.4.5 Thermodynamic inequalities

The conditions for $G_0$ to be a minimum at equilibrium impose some inequalities on the partial derivatives of the thermodynamic functions. Writing $U$ as a function of $S$ and $V$ (with $N$ and other variables fixed), we have $G_0 = U(S,V) - T_0 S + P_0 V$ so that at equilibrium

$$\frac{\partial G_0}{\partial S} = T - T_0 = 0 \qquad \frac{\partial G_0}{\partial V} = P_0 - P = 0 \tag{13.65}$$

yielding, at least formally, the equilibrium values of $S$ and $V$

$$S = S_{\text{eq}}(T_0, P_0) \qquad V = V_{\text{eq}}(T_0, P_0) \tag{13.66}$$

Near the equilibrium, we can perform a Taylor expansion of $G_0$

$$G_0(S,V) = G_0\left(S_{\text{eq}}, V_{\text{eq}}\right) + \frac{\Delta S^2}{2}\frac{\partial^2 G_0}{\partial S^2} + \frac{\Delta V^2}{2}\frac{\partial^2 G_0}{\partial V^2} + \Delta S \Delta V \frac{\partial^2 G_0}{\partial S \partial V} \tag{13.67}$$

where $\Delta S = S - S_{\text{eq}}$ and $\Delta V = V - V_{\text{eq}}$. The first order derivatives are null by definition. The second order derivatives may in turn be written as

$$\frac{\partial^2 G_0}{\partial S^2} = \left(\frac{\partial T}{\partial S}\right)_V \qquad \frac{\partial^2 G_0}{\partial V^2} = -\left(\frac{\partial P}{\partial V}\right)_S \qquad \frac{\partial^2 G_0}{\partial S \partial V} = \left(\frac{\partial T}{\partial V}\right)_S = -\left(\frac{\partial P}{\partial S}\right)_V \tag{13.68}$$

and for this to be a minimum at $(S_{\text{eq}}, V_{\text{eq}})$ the second order term must be positive for any choice of $(\Delta S, \Delta V)$. Writing $X = \Delta S / \Delta V$, this means that for any $X$, we have

$$\left(\frac{\partial T}{\partial S}\right)_V X^2 + 2\left(\frac{\partial T}{\partial V}\right)_S X - \left(\frac{\partial P}{\partial V}\right)_S > 0 \tag{13.69}$$

This quadratic form is positive if the second-degree and zeroth-degree coefficients are positive

$$\left(\frac{\partial T}{\partial S}\right)_V > 0 \qquad -\left(\frac{\partial P}{\partial V}\right)_S > 0 \tag{13.70}$$

and if the discriminant is negative [4]

$$\left[\left(\frac{\partial T}{\partial V}\right)_S\right]^2 + \left(\frac{\partial T}{\partial S}\right)_V \left(\frac{\partial P}{\partial V}\right)_S < 0 \tag{13.71}$$

4. Actually, the positivity of the second-degree coefficient and the negativity of the discriminant ensure the positivity of the zeroth-degree coefficient.

The first two conditions are equivalent to the positivity of the heat capacity at constant volume $C_v$ and of the adiabatic compressibility $\chi_S$, while the discriminant condition may be shown to imply the positivity of the isothermal compressibility $\chi_T$, a demonstration we leave to the reader,

$$C_v = T\left(\frac{\partial S}{\partial T}\right)_V > 0 \qquad \chi_S = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_S > 0 \qquad \chi_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T > 0 \tag{13.72}$$

## 13.5 Applications

### 13.5.1 Phase transition of a pure compound

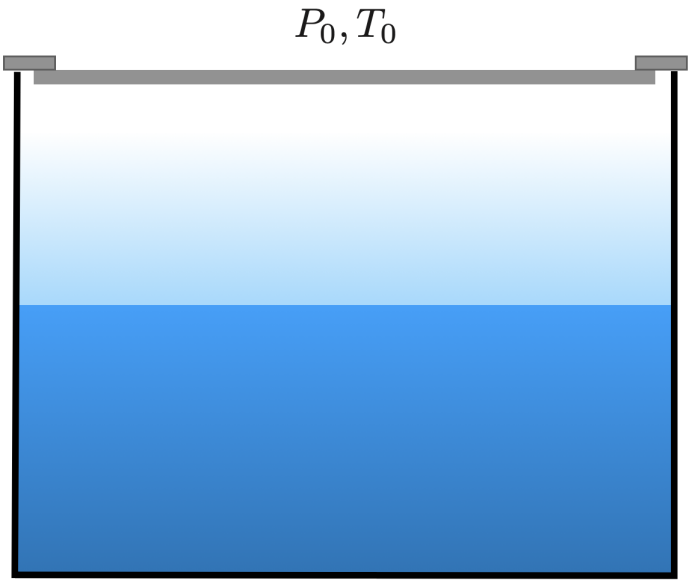



FIGURE 13.6 – A two-phase equilibrium under imposed pressure and temperature

#### Equilibrium condition

We consider a system consisting of a mass $m$ of a pure compound existing in two phases $a$ and $b$ (for instance, these could be liquid and vapor) and submitted to a fixed pressure $P_0$ and temperature $T_0$. The external parameters are then $m$, $T_0$, and $P_0$. When the two phases coexist, there is a mass $m_a$ in phase $a$ and a mass $m_b$ in phase $b$, with $m_a + m_b = m$. The internal variable here is, for instance, the mass fraction $x = m_a/(m_a + m_b)$ in phase $a$, with obviously $0 \leqslant x \leqslant 1$.

The equilibrium of the system is determined by the minimum of the Gibbs free-energy $G$. Writing $G_a(T, P)$ for the Gibbs free-energy if the mass $m$ is in phase $a$ and $G_b(T, P)$ for the Gibbs free-energy if the mass $m$ is in phase $b$, we have

$$G(x; T, P) = xG_a(T, P) + (1 - x)G_b(T, P) = G_b(T, P) + x\left[G_a(T, P) - G_b(T, P)\right] \tag{13.73}$$

The value $x_{\text{eq}}$ of $x$ at equilibrium must be such that this function reaches a minimum, so three cases are possible for a given set of temperature and pressure (Fig. 13.7) :

— If $G_a(T, P) < G_b(T, P)$ then the minimum is for $x = 1$, and the whole system is in phase $a$.

— If $G_a(T,P) > G_b(T,P)$ then the minimum is for $x = 0$, and the whole system is in phase $b$.
— If $G_a(T,P) = G_b(T,P)$ any $x$ will do and both phases coexist

In that last case, however, the condition $G_a(T,P) = G_b(T,P)$ means that there is a relation between pressure and temperature. In the case of the liquid-vapor transition, we talk of the **saturating vapor pressure** $P = P_s(T)$ when the two phases coexist.

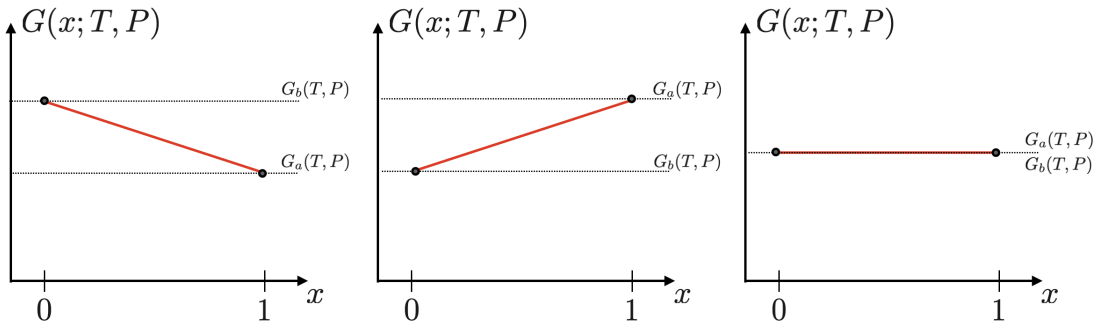



FIGURE 13.7 – $G$ as a function of the mass fraction $x$, for a given $(T,P)$, shown for the different cases discussed above.

Note that if we fix the pressure, the evolution of $G_a$ and $G_b$ as a function of temperature is related to the entropy of the two phases since

$$\left(\frac{\partial G}{\partial T}\right)_P = -S \tag{13.74}$$

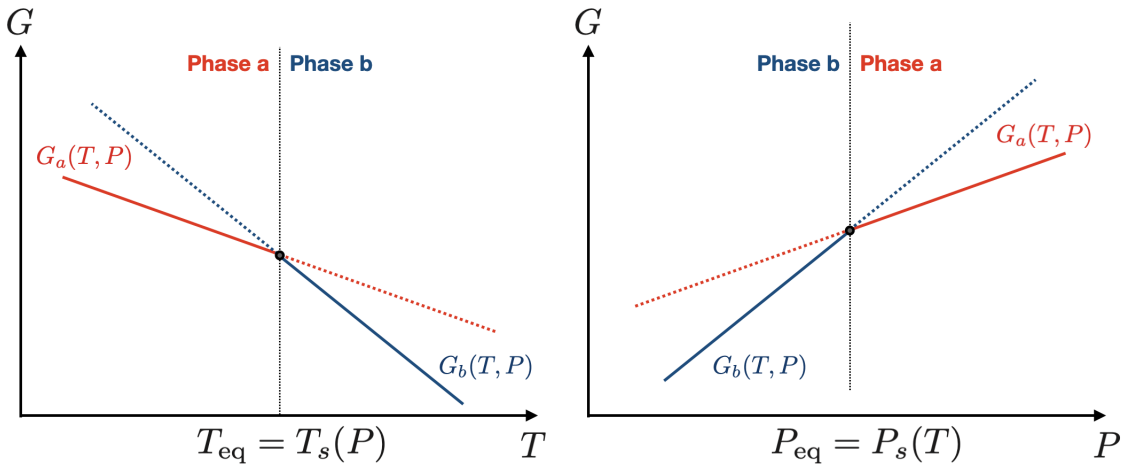



FIGURE 13.8 – $G$ as a function of $T$ for fixed $P$ (left), and as a function of $P$ for fixed $T$ (right)

So if $S_b > S_a$, we have the situation shown in Fig. 13.8, with $G_a(T,P) < G_b(T,P)$ for $T < T_{\text{eq}}(P)$ and $G_a(T,P) > G_b(T,P)$ for $T > T_{\text{eq}}(P)$, which in turn means that the system is entirely in phase $a$ below that equilibrium temperature[5] and entirely in phase $b$ above it. Similarly, we may fix the temperature, and write

$$\left(\frac{\partial G}{\partial P}\right)_T = V \tag{13.75}$$

5. which is such that $P_s\left[T_{\text{eq}}(P)\right] = P$.

If $V_b > V_a$, we conclude from Fig. 13.8 that the system is entirely in phase $b$ for $P < P_s(T)$ and entirely in phase $a$ if $P > P_s(T)$. Consequently, in the $(P,T)$ plane we can place the $a$ and $b$ phases as indicated in Fig. 13.9 and conclude that $P_s(T)$ is an increasing function of temperature. Of course, if $V_b < V_a$, we find that $P_s(T)$ is a decreasing function of temperature. This is the case for instance of the solid-liquid transition of water.

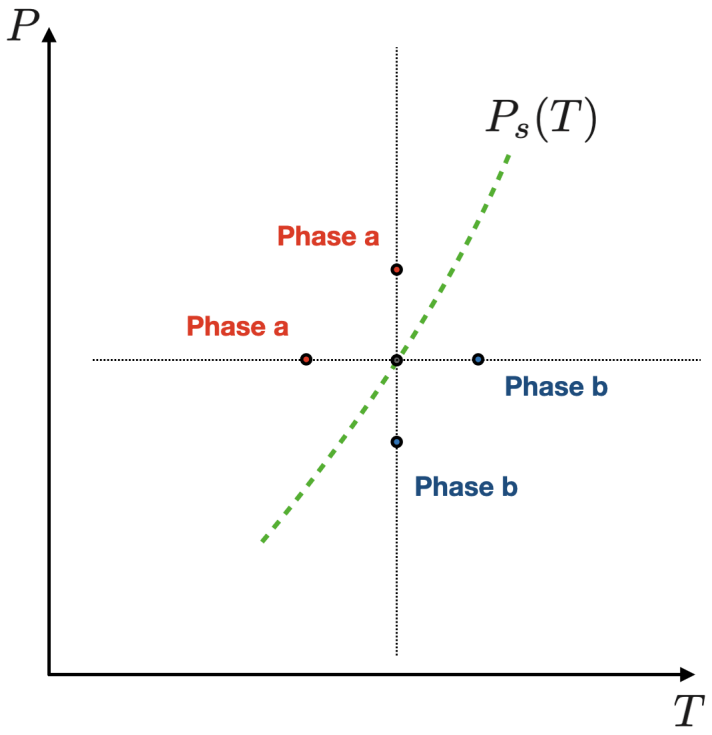



FIGURE 13.9 – Phases $a$ and $b$ in the $(P,T)$ plane, for $S_b > S_a$ and $V_b > V_a$. The green dashed curve represents the equilibrium curve between the two phases.

### Clausius-Clapeyron relation

The slope of the equilibrium curve $P_s(T)$ discussed above may be computed locally by using the condition for equilibrium $G_a(T,P) = G_b(T,P)$ at two nearby points $(T,P)$ and $(T+\mathrm{d}T, P+\mathrm{d}P)$ that are both on the curve. This implies that the variations of $G_a$ and $G_b$ are identical, i.e., $\mathrm{d}G_a = \mathrm{d}G_b$ or

$$\left(\frac{\partial G_a}{\partial T}\right)_P \mathrm{d}T + \left(\frac{\partial G_a}{\partial P}\right)_T \mathrm{d}P = \left(\frac{\partial G_b}{\partial T}\right)_P \mathrm{d}T + \left(\frac{\partial G_b}{\partial P}\right)_T \mathrm{d}P \tag{13.76}$$

This may be written using the entropy and volume, as these are the partial derivatives introduced here,

$$-S_a \mathrm{d}T + V_a \mathrm{d}P = -S_b \mathrm{d}T + V_b \mathrm{d}P \tag{13.77}$$

Rearranging this to obtain the derivative of $P = P_s(T)$ with respect to $T$, we obtain the important **Clausius-Clapeyron relation** for phase transitions,

$$\frac{\mathrm{d}P_s}{\mathrm{d}T} = \frac{S_b - S_a}{V_b - V_a} \tag{13.78}$$

where we recall that $S_a$ is the entropy of the system when it is entirely in phase $a$, $V_a$ the volume it occupies in that state, while $S_b$ and $V_b$ are the corresponding quantities when the system is entirely in phase $b$. Considering a unit mass of fluid, these are the specific quantities $s_a$, $v_a$, $s_b$, and $v_b$, and we may rewrite the above relation using the **specific latent heat** $L_v$

$$\frac{\mathrm{d}P_s}{\mathrm{d}T} = \frac{L_v}{T\Delta v} \qquad L_v = T\Delta s = T(s_b - s_a) \qquad \Delta v = v_b - v_a \tag{13.79}$$

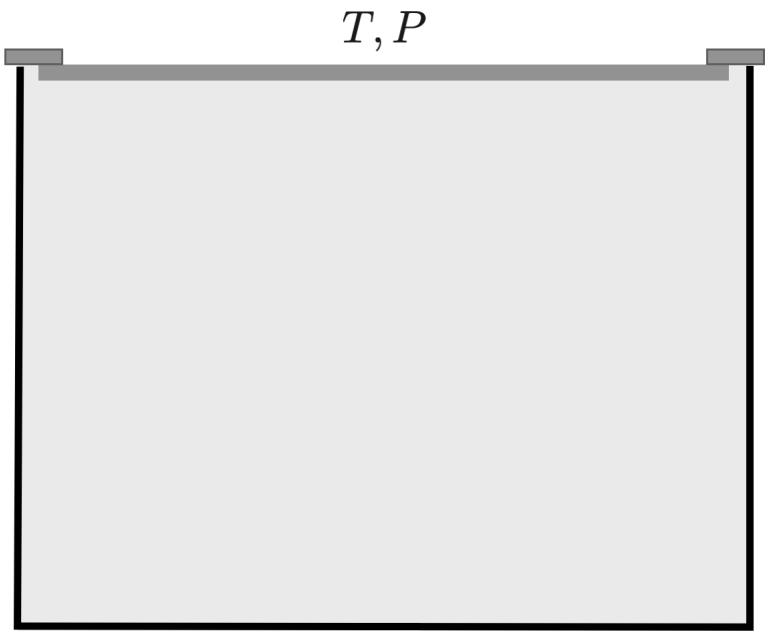



FIGURE 13.10 – Homogeneous fluid submitted to a fixed pressure and temperature

**Construction of an isothermal curve**

For a given set of external parameters $T$, $P$, and $N$, and taking the volume $V$ of the homogeneous fluid as an internal variable, the equilibrium value of the latter is found by minimising $G(V;T,P)$, yielding $V_{\mathrm{eq}}(T,P)$. Now what is the form of $G$ as a function of $V$ ? In particular, we wish to study it for a fixed $T$ (isothermal curve) and various values of the pressure. We may picture this as in Fig. 13.10, with the system in a cylinder being submitted to a user-defined pressure. If at the given temperature there is a transition, i.e., a two-phase equilibrium, this means that there is a certain pressure $P_s(T)$ for which there are two minima of $G[V;T,P_s(T)]$ corresponding to the two volumes $V_a$ and $V_b$ of the two phases. The simplest function with this property would look like the one in Fig. 13.11

Now if the pressure is not equal to $P_s(T)$ we have

$$G(V;T,P) = F(T,V) + PV = F(T,V) + P_sV + (P-P_s)\,V = G\,(V;T,P_s) + (P-P_s)\,V \tag{13.80}$$

so that compared to the case $P = P_s(T)$, the curves for $G(V;T,P)$ are in this case skewed downwards if $P < P_s(T)$ and upwards if $P > P_s(T)$, as shown in Fig. 13.12.

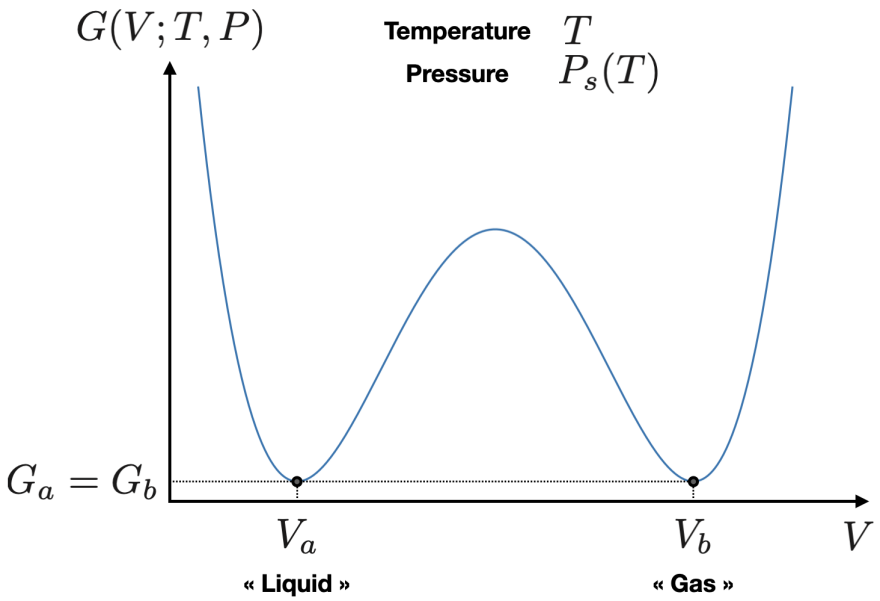



FIGURE 13.11 – $G(V;T,P)$ as a function of $V$ for a given temperature $T$ and a pressure $P = P_s(T)$.

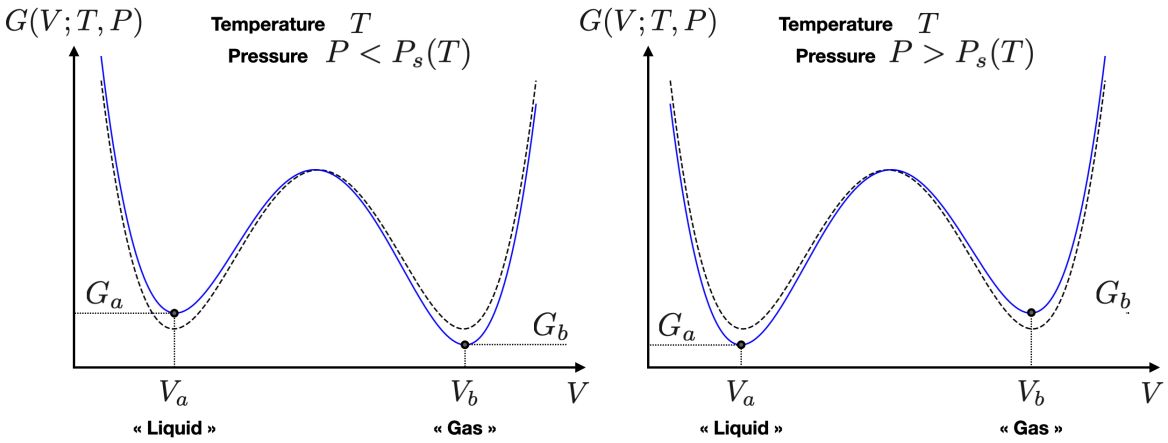



FIGURE 13.12 – $G(V;T,P)$ as a function of $V$ for a given temperature $T$ and a pressure $P \neq P_s(T)$. The case $P = P_S(T)$ is shown for comparison as the dashed line.

The two minima $G_a$ and $G_b$ are now unequal, so the higher one of the two defines a **metastable state**. Consequently, for a given temperature, and as pressure rises, the function $G(V;T,P)$ presents shapes that vary as shown in Fig. 13.13.

By reading the absolute minima of $G$ in the $(V,P)$ plane we can draw the isothermal curve, and in particular its plateau $P_S(T)$, but also the metastable branches, as shown in Fig. 13.14. Various curves thus appear :

— The **dew-point curve** is the ensemble of points for which the first drop of liquid appears, as we move from the pure vapor state.

— The **bubble-point curve** is the ensemble of points for which the first bubble of vapor appears, as we move from the pure liquid state.

— Both of these together form the **saturation curve**.

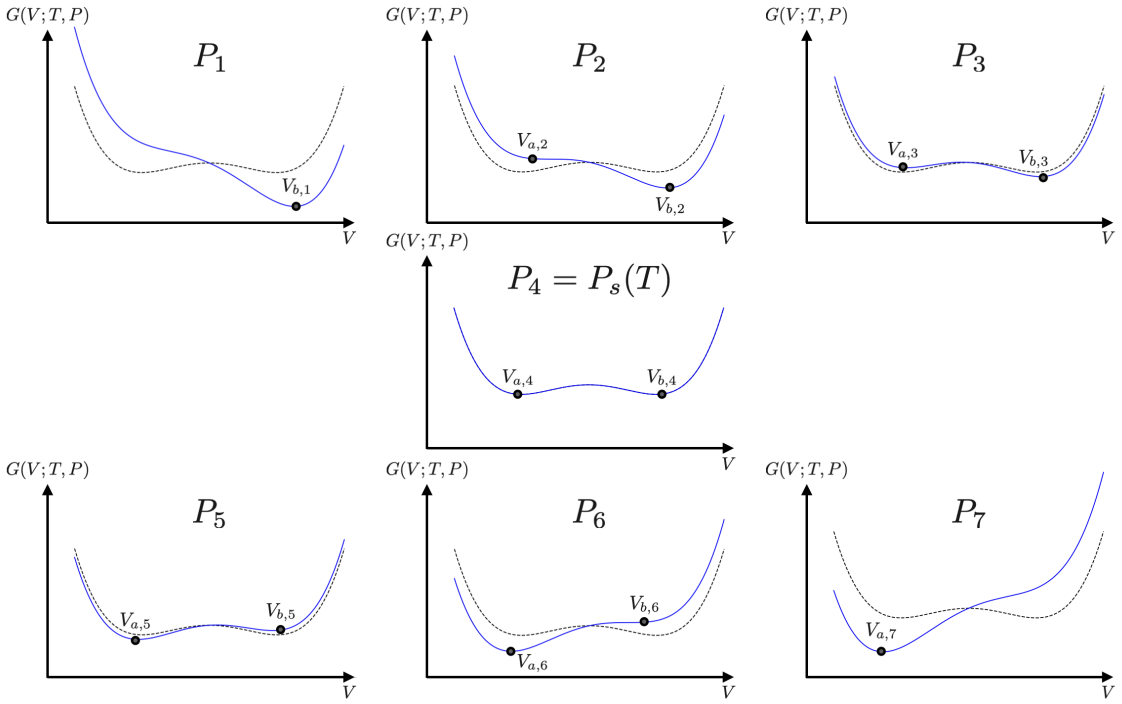



FIGURE 13.13 – $G(V;T,P)$ as a function of $V$ for a given temperature $T$ and various values of $P$ ($P_1 < P_2 < P_3 < P_4 = P_s(T) < P_5 < P_6 < P_7$).

— The **spinodal curve** is the ensemble of points where the metastable states appear or disappear. These correspond to null second-derivatives of $G$ with respect to $V$.

Between the saturation and spinodal curves, there is an equilibrium between liquid and vapor, but metastable states are possible. Below the spinodal curve, phase separation must occur, and no metastable state is allowed. The point $C$ is the **critical point** above which no clear distinction is possible between a liquid and a vapor, forming a homogeneous **fluid phase**.

### 13.5.2 Chemical reactions and the law of mass action

In a system where compounds may react chemically, i.e., modify and exchange bonds between atoms, such an evolution may occur spontaneously, with or without the action of a catalyst, leading to a **chemical equilibrium** after a sufficiently long time. Thermodynamics provides conditions for this equilibrium, independently of the path that led to it, but depending of course on the environment (temperature and pressure). To discuss these, we consider a chemical reaction written in the following symbolic form

$$\sum_i \nu_i A_i = 0 \tag{13.81}$$

where the $A_i$ stand for the various compounds at play, and the integer coefficients $\nu_i$ are either positive or negative, depending on whether the compound is a reactant or a product. We assume that this reaction occurs at constant ambient pressure and temperature, as is basically the case when working in an open vessel in the lab since the atmosphere provides the thermostat and volume reservoir necessary.

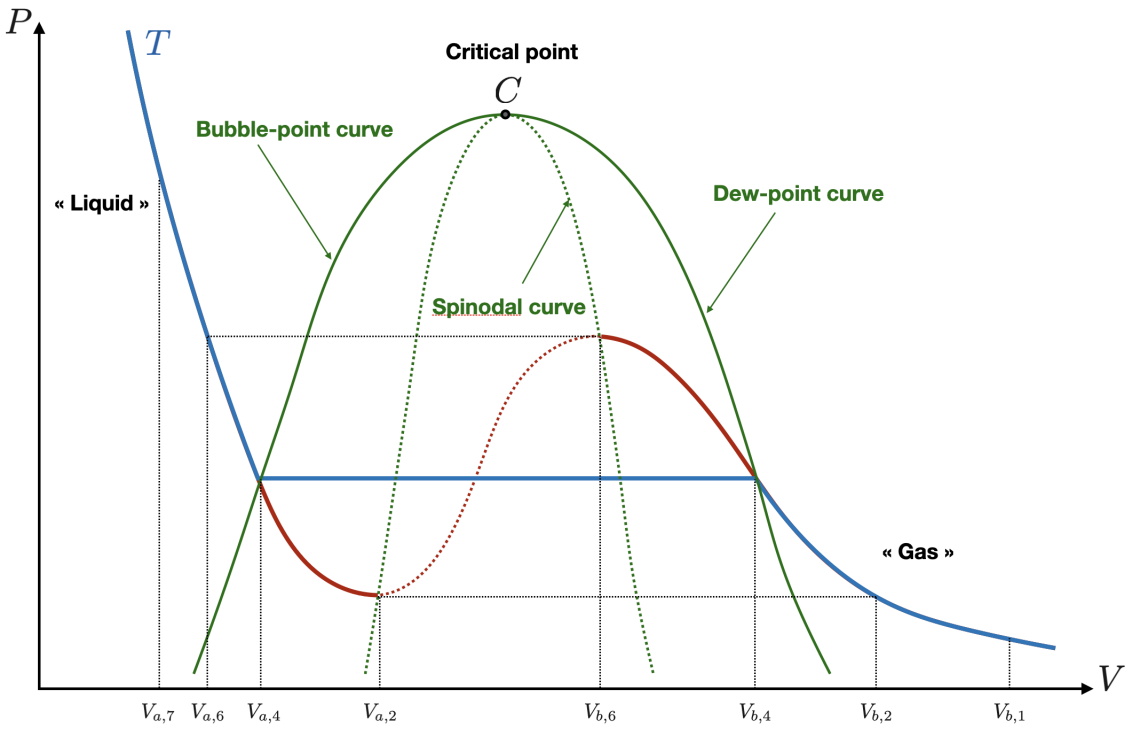



FIGURE 13.14 – Isothermal curve, metastable states, saturation curve and spinodal curve

In these conditions, as we saw, the Gibbs free-energy is minimum at equilibrium. We have

$$\mathrm{d}G = -S\mathrm{d}T + V\mathrm{d}P + \sum_i \mu_i \mathrm{d}N_i \tag{13.82}$$

where $N_i$ is the number of particles $A_i$. Now the variations of the different $N_i$ are not independent, due to the relation (13.81). Indeed, if $N_i$ varies by an amount $\nu_i$, then all other $N_j$ must vary by an amount $\nu_j$ to preserve the matter content. Isolating one species $A_i$ in particular, we then have

$$\mu_i \mathrm{d}N_i + \sum_{j\neq i} \mu_j \mathrm{d}N_j = \mathrm{d}N_i \left[\mu_i + \sum_{j\neq i} \mu_j \left(\frac{\mathrm{d}N_j}{\mathrm{d}N_i}\right)\right] = \mathrm{d}N_i \left(\mu_i + \sum_{j\neq i} \mu_j \frac{\nu_j}{\nu_i}\right) \tag{13.83}$$

so that

$$\mathrm{d}G = -S\mathrm{d}T + V\mathrm{d}P + \mathrm{d}N_i \left(\mu_i + \sum_{j\neq i} \mu_j \frac{\nu_j}{\nu_i}\right) = -S\mathrm{d}T + V\mathrm{d}P + \frac{\mathrm{d}N_i}{\nu_i} \sum_j \nu_j \mu_j \tag{13.84}$$

At $T$ and $P$ fixed and at equilibrium, we have

$$\left(\frac{\partial G}{\partial N_i}\right)_{T,P} = 0 \tag{13.85}$$

which yields the **condition for chemical equilibrium**

$$\sum_i \nu_i \mu_i = 0 \tag{13.86}$$

It may be used to derive an important result for gas-phase reactions. For that, we note that the chemical potential $\mu_i$ of a gas in a mixture is given by

$$\mu_i(T, P) = k_B T \ln P_i + \chi_i(T) \tag{13.87}$$

where $P_i$ is the partial pressure of this gas in the mixture, that is the pressure it would have if it were the sole compound in the full volume

$$P_i = \frac{N_i k_B T}{V} = \frac{N_i}{N} P \qquad\qquad N = \sum_i N_i \tag{13.88}$$

and $\chi_i(T)$ is a function of the temperature that does not need to be specified. At equilibrium, these partial pressures take values $P_{0,i}$ and the condition above yields

$$\sum_i \nu_i \left[k_B T \ln P_{0,i} + \chi_i(T)\right] = 0 \tag{13.89}$$

or, rearranging to isolate the partial pressures, we obtain the **law of mass action**

$$\prod_i P_{0,i}^{\nu_i} = K_P(T) \qquad\qquad K_P(T) = \exp\left(-\frac{1}{k_B T}\sum_i \nu_i \chi_i\right) \tag{13.90}$$

where $K_P(T)$ is the **chemical equilibrium constant**, and is characteristic of the equilibrium, and independent of the initial quantities of the various compounds. The law of mass action may be applied to solutions instead of gas-phase reactions, in which case the concentrations $c_i = N_i/N$ of the solutes replace the partial pressures. The law of mass action in this case reads

$$\prod_i c_{0,i}^{\nu_i} = K(P, T) \tag{13.91}$$

but the constant in this case also depends on the total pressure in a manner that is not analytically simple. We leave it to the reader to determine how $K_P(T)$ depends on $P$ in the gas-phase case.

### 13.5.3 Ionization equilibrium : Saha's equation

An important case of an equilibrium in astrophysics is that of the ionization equilibrium at a given temperature $T$. We consider the equilibrium described by Fig. 13.15, i.e.,

$$A_{r,i} \leftrightharpoons A_{r+1,j} + e^-(p) \tag{13.92}$$

where $A_{r,i}$ stands for state $i$ of ion $A_r$, and $A_{r+1,j}$ for state $j$ of ion $A_{r+1}$. We write $\alpha$ and $\beta$ for the states on either side of the chemical equation, i.e. $\alpha = \{A_{r,i}\}$ and $\beta = \{A_{r+1,j} + e^-(p)\}$. We thus make it clear that the energy $E_e$ of the released electron is a function of its momentum $p$. In the non-relativisitc limit, which we suppose is the case here, we have

$$E_e = \frac{p^2}{2m_e} \tag{13.93}$$

The Boltzmann distribution gives the ratio of probabilities to find the system in either of the two states

$$\frac{P_\beta}{P_\alpha} = \frac{g_\beta}{g_\alpha} \exp\left(-\frac{E_\beta - E_\alpha}{k_B T}\right) \tag{13.94}$$

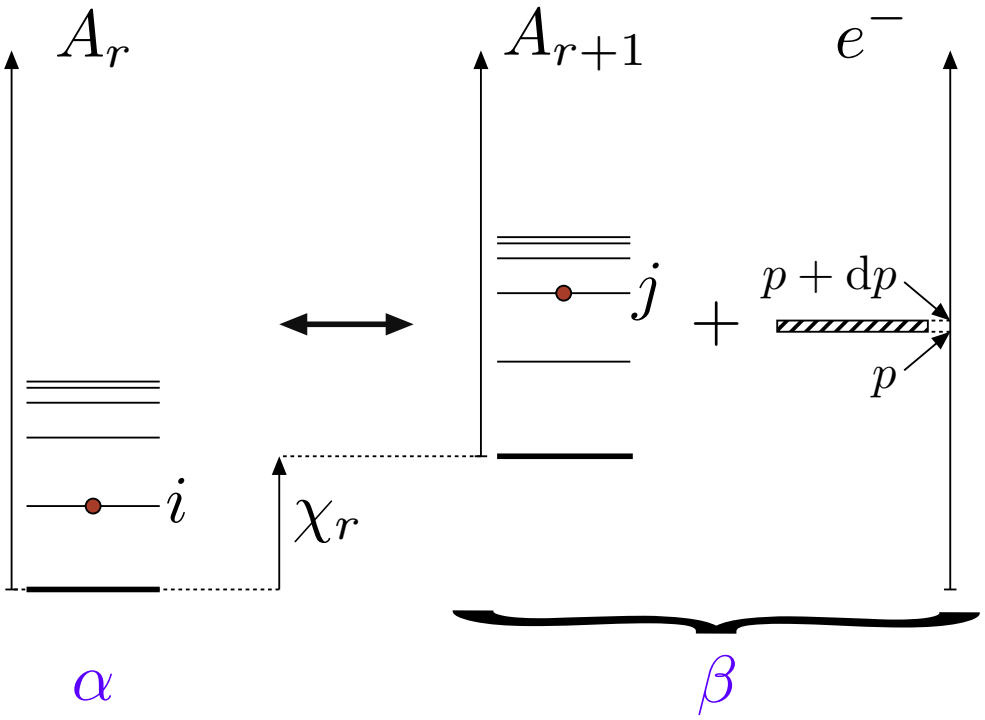



FIGURE 13.15 – Ionization equilibrium between states $r$ and $r+1$ times ionized.

Now the energies and statistical weights of these two states are, respectively,

$$E_\alpha = \mathcal{E}_{r,i} \qquad E_\beta = \mathcal{E}_{r+1,j} + \chi_r + \frac{p^2}{2m_e} \quad g_\alpha = g_{r,i} \qquad g_\beta = g_{r+1,j} g_e(p) \tag{13.95}$$

In the last equation, $g_e(p)$ represents the statistical weight of the states of the released electron. To compute this, we recall that this is just the number of microstates for an electron with momentum $p$, to within $\mathrm{d}p$. The classical computation of the density of states for a gas of $N_e$ free electrons in a box of volume $V$ shows that the number of electrons whose energy lies between $E$ and $E + \mathrm{d}E$ is

$$N(E)\mathrm{d}E = \frac{V}{2\pi^2}\left(\frac{2m_e}{\hbar^2}\right)^{3/2}\sqrt{E}\mathrm{d}E \tag{13.96}$$

which can be rearranged to exhibit $g_e(p)$, i.e., with $N(E)\mathrm{d}E = N_e g_e(p)\mathrm{d}p$ we have

$$g_e(p) = \frac{V}{N_e}\frac{p^2}{\pi^2\hbar^3} = \frac{8\pi p^2 \mathrm{d}p}{n_e h^3} \tag{13.97}$$

where $n_e = N_e/V$ is the volume density of free electrons. Consequently, the ratio of the two ionic populations is obtained by integrating over the possible momenta of the released electron, yielding

$$\frac{n_{r+1,j}}{n_{r,i}} = \frac{g_{r+1,j}}{g_{r,i}}\frac{2(2\pi m_e k_B T)^{3/2}}{n_e h^3}\exp\left(-\frac{\mathcal{E}_{r+1,j} + \chi_r - \mathcal{E}_{r,i}}{k_B T}\right). \tag{13.98}$$

Now, the Boltzmann distribution gives the populations of the various states for each ion, i.e.,

$$\frac{n_{r,i}}{n_r} = \frac{g_{r,i}}{Z_r(T)}\exp\left(-\frac{\mathcal{E}_{r,i}}{kT}\right) \qquad \frac{n_{r+1,j}}{n_{r+1}} = \frac{g_{r+1,j}}{Z_{r+1}(T)}\exp\left(-\frac{\mathcal{E}_{r+1,j}}{kT}\right) \tag{13.99}$$

where $Z_r$ and $Z_{r+1}$ are the canonical partition functions for the two ionic species, defined by

$$Z_r(T) = \sum_{j \geqslant 0} g_{r,j} \exp\left(-\frac{\mathcal{E}_{r,j}}{kT}\right) \qquad Z_{r+1}(T) = \sum_{j \geqslant 0} g_{r+1,j} \exp\left(-\frac{\mathcal{E}_{r+1,j}}{kT}\right) \tag{13.100}$$

This allows to replace the factors depending on the specific state in the previous equation by global factors $n_r/Z_r$ and $n_{r+1}/Z_{r+1}$, yielding eventually **Saha's equation** in the form

$$\frac{n_{r+1} n_e}{n_r} = \frac{2 Z_{r+1}}{Z_r} \frac{(2\pi m_e kT)^{3/2}}{h^3} \exp\left(-\frac{\chi_r}{kT}\right). \tag{13.101}$$

It has applications for instance to the determination of the ionization state of hot gaseous nebulæ.

# 14

# Evolution towards equilibrium in statistical physics

In the previous chapter, we discussed the evolution of an out-of-equilibrium system in the context of classical thermodynamics, where, depending on the experimental conditions, it is characterized by an increase of the entropy $S$ or a decrease of the Helmholtz free-energy $F$, the Gibbs free-energy $G$ or the grand-canonical potential $J$. In fact, the latter three are consequences of the former. So it is natural to wonder whether this increase of $S$ is itself a consequence of a more fundamental result in the framework of statistical physics. Going back to the fundamental postulate we introduced in Chapter 2, according to which $p_\ell = 1/\Omega$, we see that it deals with states of macroscopic equilibrium, where the system may be in any of a large number of microstates $|\ell\rangle$ with probabilities $p_\ell$ that no dot depend on time. Statistical physics is nevertheless able to tackle systems that are out-of-equilibrium, by considering time-dependent probability distributions $\{p_\ell(t)\}$, ruled by the **master equation**.

## 14.1 The master equation

### 14.1.1 Transitions between microstates

In a perfectly isolated system, the Hamiltonian $H$ could be perfectly known, and time-independent. Its eigenstates $|\ell\rangle$ would be stationary, with eigenvalues $E_\ell$ such that

$$H\,|\ell\rangle = E_\ell\,|\ell\rangle \tag{14.1}$$

In reality, however, there is some uncertainty in the Hamiltonian, which reads $H + \delta H$, where the uncertainty in $H$ must be of order $\delta E$, the experimental uncertainty in the energy of the system. Remember that the number of accessible states $\Omega(E;\delta E)$ actually takes into account this uncertainty, as the notation underlines. This leads to constant changes in the system from one accessible microstate $|\ell\rangle$ to another $|m\rangle$. Indeed, there are always forces originating from outside the system that are too weak to be taken into account explicitly in the Hamiltonian but nevertheless induce transitions from one quantum state to another[1], as shown schematically in Fig. 14.1.

These transitions are characterized by a **transition probability** $\mathcal{P}_{m\ell}(t_0, t_0 + \mathrm{d}t)$ that the system, being in state $|\ell\rangle$ at time $t_0$, transitions to find itself in state $|m\rangle$ at the later time $t_0 + \mathrm{d}t$. Of course $\mathcal{P}_{m\ell}(t_0, t_0 + \mathrm{d}t) \to \delta_{m\ell}$ in the limit $\mathrm{d}t \to 0$, but the characteristic transition time $\tau$ is much smaller than the timescale $\Delta t$ over which the sytem is observed, making the transition process a very rapid one. The memory loss means that $t_0$ is meaningless. So let us consider a time interval $\mathrm{d}t$ such that

$$\tau \ll \mathrm{d}t \ll \Delta t \tag{14.2}$$

1. See the discussion at the end of this chapter.

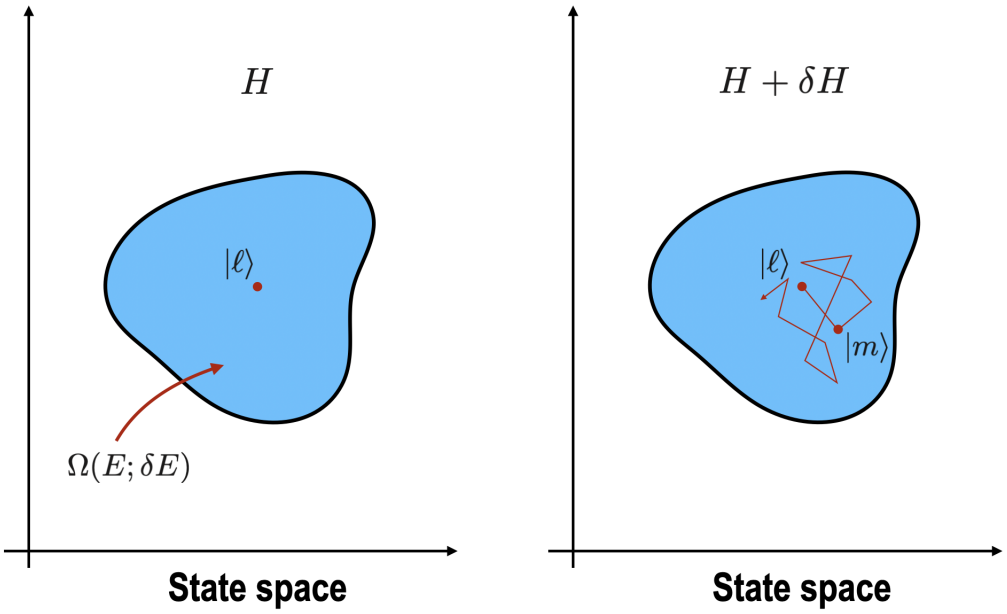



FIGURE 14.1 – Stationary states in a perfectly known Hamiltonian (*left*) and transitions between accessible microstates in the presence of (unavoidable) perturbations of the Hamiltonian (*right*).

We then have, for $|m\rangle \neq |\ell\rangle$

$$\mathcal{P}_{m\ell}(t_0, t_0 + \mathrm{d}t) = a_{m\ell}\mathrm{d}t \tag{14.3}$$

where $a_{m\ell}$ is the transition probability per unit time[2]. It does not depend on time fr a system in macroscopic equilibrium, and is of course positive, $a_{m\ell} \geqslant 0$ for $|m\rangle \neq |\ell\rangle$. Note that we can define an $a_{\ell\ell} < 0$ by writing

$$\mathcal{P}_{\ell\ell}(t_0, t_0 + \mathrm{d}t) = 1 + a_{\ell\ell}\mathrm{d}t \tag{14.4}$$

such that the probability to stay in state $|\ell\rangle$ is one in the limit $\mathrm{d}t \to 0$ and is less than one at finite times.

### 14.1.2 The master equation for an isolated system

How do the probabilities $p_\ell$ of microstates $|\ell\rangle$ evolve with time for an isolated system ? This is a simple matter of counting transitions, between $t$ and $t + \mathrm{d}t$, into and out the microstate $|\ell\rangle$. This gives a change

$$\mathrm{d}p_\ell = \sum_m a_{\ell m} p_m(t)\mathrm{d}t - \sum_m a_{m\ell} p_\ell(t)\mathrm{d}t \tag{14.5}$$

where the first sum gives the probability increase from transitions into state $|\ell\rangle$ and the second sum gives the probability decrease from transitions out of state $|\ell\rangle$. This gives the **master equation**

$$\frac{\mathrm{d}p_\ell}{\mathrm{d}t} = \sum_m \left[a_{\ell m} p_m(t) - a_{m\ell} p_\ell(t)\right] \tag{14.6}$$

Note that the sums can include the terms for $|m\rangle = |\ell\rangle$, since these cancel out. To complete this into an **evolution postulate**, we add two hypotheses[3] :

2. This should remind the reader of the Einstein $A_{ul}$ coefficient of spontaneous emission seen in Chapter 7.
3. The master equation and these hypotheses are in fact results from more fundamental ideas from quantum mechanics.

— The transition matrix is **symmetric**, i.e., $a_{\ell m} = a_{m\ell}$
— There is no transition ($a_{\ell m} = 0$) between states whose energy difference $|E_\ell - E_m|$ is larger than the experimental uncertainty $\delta E$.

One should note that the master equation is in fact a system of coupled differential equations of order one in time. Therefore, it is not invariant in time reversal $t \to -t$, unlike the equations of motion at the microscopic level [4]. This is a very important property to understand **irreversibility** in the evolution of macroscopic systems.

### 14.1.3 The master equation for a system in contact with a thermostat

From the master equation for an isolated system, it is possible to deduce what form it takes for a system in contact with a thermostat, by using the same process as in Chapter 3. We write $\mathcal{S}$ for the system under study and $\mathcal{T}$ for the thermostat. The union system $\mathcal{S} \bigcup \mathcal{T}$ is assumed to be isolated, with a weak coupling between $\mathcal{S}$ and $\mathcal{T}$, and therefore its probability distribution $p_{\ell,L}(t)$ obeys the master equation written above, with $|\ell, L\rangle$ the microstate of the union system, made up of a state $|\ell\rangle$ of $\mathcal{S}$ and a state $|L\rangle$ of $\mathcal{T}$. Writing $a_{\ell,L;m,M}$ for the transition probabilities per unit time in the union system, we have

$$\frac{\mathrm{d}p_{\ell,L}}{\mathrm{d}t} = \sum_m \sum_M \left[a_{\ell,L;m,M} p_{m,M} - a_{m,M;\ell,L} p_{\ell,L}\right] \tag{14.7}$$

Of course, we are interested in the probability distribution of the sole system $\mathcal{S}$, for which we have

$$p_\ell = \sum_L p_{\ell,L} = \sum_L p_\ell p_L \left(E_0 - E_\ell\right) \tag{14.8}$$

where $E_0 = E_\ell + E_L$ is the total energy of the union system, that is conserved. From this, we have

$$\frac{\mathrm{d}p_\ell}{\mathrm{d}t} = \sum_L \frac{\mathrm{d}p_{\ell,L}}{\mathrm{d}t} = \sum_L \sum_m \sum_M \left[a_{\ell,L;m,M} p_m p_M \left(E_0 - E_m\right) - a_{m,M;\ell,L} p_\ell p_L \left(E_0 - E_\ell\right)\right] \tag{14.9}$$

We switch sums and extract $p_m$ and $p_\ell$ from the remaining sums over $L$ and $M$

$$\frac{\mathrm{d}p_\ell}{\mathrm{d}t} = \sum_m \left\{ p_m \sum_L \sum_M \left[a_{\ell,L;m,M} p_M \left(E_0 - E_m\right)\right] - p_\ell \sum_L \sum_M \left[a_{m,M;\ell,L} p_L \left(E_0 - E_\ell\right)\right] \right\} \tag{14.10}$$

By introducing the following coefficients

$$a^T_{\ell m} = \sum_L \sum_M a_{\ell,L;m,M} p_M \left(E_0 - E_m\right) \qquad a^T_{m\ell} = \sum_L \sum_M a_{m,M;\ell,L} p_L \left(E_0 - E_\ell\right) \tag{14.11}$$

we recover a master equation for the probability distribution of $\mathcal{S}$ as [5]

$$\frac{\mathrm{d}p_\ell}{\mathrm{d}t} = \sum_m \left[a^T_{\ell m} p_m - a^T_{m\ell} p_\ell\right] \tag{14.12}$$

The coefficients $a^T_{\ell m}$ are time-independent and positive for $|\ell\rangle \neq |m\rangle$, but they are not symmetric. This property is replaced by the following relation

$$a^T_{\ell m} \exp\left(-\frac{E_m}{k_B T}\right) = a^T_{m\ell} \exp\left(-\frac{E_\ell}{k_B T}\right) \tag{14.13}$$

4. The case of a classical equation of motion, that is second-order in time, is clearly time-revertible. The case of the Schrödinger equation is less clear, but the fact that it involves complex-valued wave functions ensures its time-reversal invariance as well.

5. It is direct to show that the definitions of $a^T_{\ell m}$ and $a^T_{m\ell}$ above are consistent with each other.

which stems from the fact that these transition probabilities may link two microstates $|\ell\rangle$ and $|m\rangle$ with significant energy differences, the thermostat absorbing or furnishing the heat necessary involved in this transition. Indeed, by expressing the microcanonical probability distribution of the thermostat as

$$P_L(E_0 - E_\ell) = A \exp\left(\frac{E_\ell}{k_B T}\right) \tag{14.14}$$

if the energy of the microstate $|L\rangle$ is compatible with $E_0 - E_\ell$ to within the experimental uncertainty, and $P_L(E_0 - E_\ell) = 0$ otherwise, we have

$$a^T_{\ell m} = \sum_L \sum_M a_{\ell,L;m,M} A \exp\left(-\frac{E_0 - E_m}{k_B T}\right) \qquad a^T_{m\ell} = \sum_L \sum_M a_{m,M;\ell,L} A \exp\left(-\frac{E_0 - E_\ell}{k_B T}\right) \tag{14.15}$$

which gives the result, using the symmetry of the $a_{\ell,L;m,M}$,

$$a^T_{\ell m} \exp\left(-\frac{E_m}{k_B T}\right) = \sum_L \sum_M a_{\ell,L;m,M} A \exp\left(-\frac{E_0}{k_B T}\right) = a^T_{m\ell} \exp\left(-\frac{E_\ell}{k_B T}\right) \tag{14.16}$$

Assuming for instance that $E_\ell > E_m$, the ratio of the two transition probabilities per unit time

$$\frac{a^T_{\ell m}}{a^T_{m\ell}} = \exp\left(-\frac{E_\ell - E_m}{k_B T}\right) < 1 \tag{14.17}$$

shows that there are more transitions from the upper level $|\ell\rangle$ to the lower one $|m\rangle$ than the opposite.

## 14.2 Consequences of the master equation

### 14.2.1 Detailed balance at equilibrium

#### Equilibrium solutions

At equilibrium, the probabilities do not depend on time. To make this explicit, we write $p^{(e)}_\ell$ for the probabilities at equilibrium. We hence have, for any state $|\ell\rangle$,

$$\frac{\mathrm{d}p^{(e)}_\ell}{\mathrm{d}t} = \sum_m \left[a_{\ell m} p^{(e)}_m - a_{m\ell} p^{(e)}_\ell\right] = 0 \tag{14.18}$$

which does not mean that the system stays in a given microstate at all times, but that, given an ensemble of systems, there are as many of these systems that transition into state $|\ell\rangle$ as there are that transition out of it, in a given unit of time. To find the equilibrium probabilities, one has to solve a linear system of many (countable) equations, subject to the constraints

$$\forall\, |\ell\rangle \;\; p^{(e)}_\ell \geqslant 0 \qquad \sum_{|\ell\rangle} p^{(e)}_\ell = 1 \tag{14.19}$$

We shall now compute these equilibrium solutions in the two cases of an isolated system and of a system in contact with a thermostat.

#### The case of an isolated system : recovering the microcanonical distribution

It is straightforward to show that the microcanonical distribution is an equilibrium solution of the master equation. Indeed, let us take

$$p_\ell = \frac{1}{\Omega(E)} \tag{14.20}$$

for any accessible state $|\ell\rangle$ among the $\Omega(E)$ that exist, and $p_\ell = 0$ otherwise. We consider the sum

$$\sum_m \left[a_{\ell m} p_m - a_{m\ell} p_\ell\right] \tag{14.21}$$

Assume first that $|\ell\rangle$ is not an accessible state, then the second part of the sum is zero by virtue of $p_\ell = 0$. In the first, only the terms for which $|m\rangle$ is an accessible state have non-zero $p_m$. However, in that case the transition probabilities are $a_{\ell m} = 0$, so that the whole sum is null. In the case that $|\ell\rangle$ is an accessible state, the only terms that survive in both parts of the sum are those that connect $|\ell\rangle$ to another accessible state $|m\rangle$, and the symmetry property $a_{\ell m} = a_{m\ell}$ combined with the equality $p_\ell = p_m$ for these accessible states ensures that the whole sum is also null. Interestingly, in that latter case, the sum is null because of the cancellation of individual terms

$$a_{\ell m} p_m = a_{m\ell} p_\ell \tag{14.22}$$

for all accessible states $|\ell\rangle$ and $|m\rangle$, even though equilibrium only requires that the sum (14.21) be null. This very specific condition (14.22) is called **detailed balance**, in which case not only are the probabilities to leave state $|\ell\rangle$ and to enter it equal during a given time interval $\mathrm{d}t$, but they are also equal when considering a specific ingoing and outgoing state $|m\rangle$. Finally, it is possible to show that the microcanonical distribution is the only equilibrium solution in the case of an isolated system, as we shall see later.

#### The case of a system in contact with a thermostat : recovering the canonical distribution

In the case of a system in contact with a thermostat, it is straightforward to show that the canonical distribution is an equilibrium solution. Indeed, taking

$$p_\ell = \frac{1}{Z} \exp\left(-\frac{E_\ell}{k_B T}\right) \tag{14.23}$$

we have, using the relationship (14.13) between $a^T_{\ell m}$ and $a^T_{m\ell}$,

$$a^T_{\ell m} p_m = \frac{a^T_{\ell m}}{Z} \exp\left(-\frac{E_m}{k_B T}\right) = \frac{a^T_{m\ell}}{Z} \exp\left(-\frac{E_\ell - E_m}{k_B T}\right) \exp\left(-\frac{E_m}{k_B T}\right) = a^T_{m\ell} p_\ell \tag{14.24}$$

ensuring that this probability distribution is an equilibrium one in this case

$$\frac{\mathrm{d}p_\ell}{\mathrm{d}t} = \sum_m \left[a^T_{\ell m} p_m - a^T_{m\ell} p_\ell\right] = 0 \tag{14.25}$$

with a detailed balance property. This equilibrium probability distribution is the only possible one, as we shall see later.

### 14.2.2 Spontaneous evolution

#### Evolution towards equilibrium

The master equation is a set of first-order time-differential equations. As such, it suffices to give a distribution $\{p_\ell(t_0)\}$ at a time $t_0$ to determine the full time-evolution $\{p_\ell(t)\}$ of the distribution. It may be shown that

$$\lim_{t\to\infty} p_\ell(t) = p_\ell^{(e)} \tag{14.26}$$

meaning that in the limit of large times, any probability distribution will tend to the equilibrium distribution (microcanonical in the case of an isolated system, canonical in the case of a system in contact with a thermostat, etc.). Any macroscopic system will, given enough time, reach its macroscopic equilibrium

state. This involves discussing **relaxation times** $\tau_r$, that appear in many systems as an exponential time scale in the approach to equilibrium

$$\left| p_\ell(t) - p_\ell^{(e)} \right| \sim c_\ell e^{-t/\tau_r} \tag{14.27}$$

Few things can be said about these times, as their computation depends heavily on the details of the system [6], but it is clear that they increase dramatically with the size of the system, as equilibrium is first reached locally before extending to more distant parts. They can vary by orders of magnitude (e.g., $10^{-12}\,\mathrm{s}$ for the rearrangement of electrons in a metal, hundreds of years for a glass-crystal transition...) and there can be several relaxation times in a system, referring to the various equilibria at stake (particle and heat diffusion, for instance). Finally, one should consider how $\tau_r$ relates to the experimental timescale $\Delta t$, for if $\tau_r \ll \Delta t$ one can readily use the formalism of equilibrium statistical physics. The other extreme $\tau_r \gg \Delta t$ is also relatively simple as the system can be thought of as being in a **metastable state**. Only the intermediate case $\tau_r \sim \Delta t$ requires a full time-dependent treatment based on the master equation.

### The $H$ theorem for an isolated system

The evolution of the system is dictated by the fact that it should bring it to equilibrium. **Boltzmann's $H$ theorem** allows to characterize this evolution. Consider an isolated system that is out of equilibrium. Its statistical entropy is given by the general formula

$$S = -k_B \sum_{|\ell\rangle} p_\ell \ln p_\ell \tag{14.28}$$

that depends on time since the distribution of probabilities does. Let us write its time-derivative as

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -k_B \left[ \sum_{|\ell\rangle} \frac{\mathrm{d}p_\ell}{\mathrm{d}t} \ln p_\ell + \sum_{|\ell\rangle} p_\ell \frac{1}{p_\ell} \frac{\mathrm{d}p_\ell}{\mathrm{d}t} \right] = -k_B \left[ \sum_{|\ell\rangle} \frac{\mathrm{d}p_\ell}{\mathrm{d}t} \ln p_\ell + \sum_{|\ell\rangle} \frac{\mathrm{d}p_\ell}{\mathrm{d}t} \right] \tag{14.29}$$

The second sum is null by virtue of the normalization, and in the first we use the master equation

$$\frac{\mathrm{d}S}{\mathrm{d}t} = -k_B \sum_{|\ell\rangle} \sum_{|m\rangle} \left( a_{\ell m} p_m - a_{m\ell} p_\ell \right) \ln p_\ell \tag{14.30}$$

Using the symmetry of coefficients, we may write this in either of the following two forms

$$\frac{\mathrm{d}S}{\mathrm{d}t} = k_B \sum_{|\ell\rangle} \sum_{|m\rangle} a_{\ell m} \left( p_\ell - p_m \right) \ln p_\ell \qquad \frac{\mathrm{d}S}{\mathrm{d}t} = k_B \sum_{|\ell\rangle} \sum_{|m\rangle} a_{\ell m} \left( p_m - p_\ell \right) \ln p_m \tag{14.31}$$

We can also write it as the half-sum of these two expressions

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{k_B}{2} \sum_{|\ell\rangle} \sum_{|m\rangle} a_{\ell m} \left( p_\ell - p_m \right) \left( \ln p_\ell - \ln p_m \right) \tag{14.32}$$

In this expression, terms are either positive or null, since $x \mapsto \ln x$ is an increasing function, so that if $p_\ell - p_m > 0$ then $\ln p_\ell - \ln p_m > 0$ and conversely. Consequently,

$$\frac{\mathrm{d}S}{\mathrm{d}t} \geqslant 0 \tag{14.33}$$

6. These are determined by the values of the coefficients $a_{\ell m}$, needed to solve the differential equation system.

which is a demonstration of the second principle of thermodynamics using only the master equation and the symmetry of its coefficients. Incidentally, the equilibrium is reached when this time-derivative is zero, implying that each of the terms in (14.32) is null. For all pairs $(|\ell\rangle, |m\rangle)$ of accessible states[7], this implies $p_\ell = p_m$, and we thus recover the microcanonical distribution.

Summarizing, for an isolated system that is initially out of equilibrium, the spontaneous evolution is in the sense of an increase of the statistical entropy, which reaches a maximum at equilibrium,at which point the probability distribution of the accessible states is microcanonical. In that sense, the fundamental postulate of 2.2.2 is in fact a consequence of the evolution postulate.

### The $H$ theorem for a system in contact with a thermostat

In the case of a system $\mathcal{S}$ in contact with a thermostat $\mathcal{T}$, the $H$ theorem applies to the (weakly-coupled) union $\mathcal{S} \bigcup \mathcal{T}$. The entropy of the system $\mathcal{S}$ may decrease if that of $\mathcal{T}$ increases more. In that case, let us consider now its Helmholtz free energy $F = \langle E \rangle - TS$, where

$$\langle E(t) \rangle = \sum_{|\ell\rangle} p_\ell(t) E_\ell \tag{14.34}$$

is the (time-dependent) average energy of the system $\mathcal{S}$. The Helmholtz free-energy is then

$$F = \langle E \rangle - TS = \sum_{|\ell\rangle} p_\ell E_\ell + T k_B \sum_\ell p_\ell \ln p_\ell = \sum_{|\ell\rangle} p_\ell \left( E_\ell + k_B T \ln p_\ell \right) \tag{14.35}$$

Let us now consider its time evolution,

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t} \left[ \sum_{|\ell\rangle} p_\ell \left( E_\ell + k_B T \ln p_\ell \right) \right] = \sum_{|\ell\rangle} \left[ \frac{\mathrm{d}p_\ell}{\mathrm{d}t} \left( E_\ell + k_B T \ln p_\ell \right) \right] + k_B T \sum_{|\ell\rangle} \frac{\mathrm{d}p_\ell}{\mathrm{d}t} \tag{14.36}$$

The last sum is null by virtue of the normalization, and we have, by the master equation in this case

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \sum_{|\ell\rangle} \left[ \sum_{|m\rangle} \left[ a^T_{\ell m} p_m - a^T_{m\ell} p_\ell \right] \left( E_\ell + k_B T \ln p_\ell \right) \right] \tag{14.37}$$

Introducing the alternate coefficients

$$\tilde{a}^T_{m\ell} = a^T_{m\ell} \exp\left( -\frac{E_\ell}{k_B T} \right) \tag{14.38}$$

that are now symmetric by virtue of (14.13), we have

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \sum_{|\ell\rangle} \sum_{|m\rangle} \left[ \tilde{a}^T_{\ell m} \exp\left( \frac{E_m}{k_B T} \right) p_m - \tilde{a}^T_{m\ell} \exp\left( \frac{E_\ell}{k_B T} \right) p_\ell \right] k_B T \ln \left[ p_\ell \exp\left( \frac{E_\ell}{k_B T} \right) \right] \tag{14.39}$$

By posing

$$\tilde{p}_\ell = p_\ell \exp\left( \frac{E_\ell}{k_B T} \right) \tag{14.40}$$

we have

$$\frac{\mathrm{d}F}{\mathrm{d}t} = k_B T \sum_{|\ell\rangle} \sum_{|m\rangle} \left[ \tilde{a}^T_{\ell m} \tilde{p}_m - \tilde{a}^T_{m\ell} \tilde{p}_\ell \right] \ln \tilde{p}_\ell \tag{14.41}$$

7. A necessary condition for $a_{\ell m}$ not to be zero.

Following the same ideas as for the case of an isolated system, this can be written

$$\frac{\mathrm{d}F}{\mathrm{d}t} = -\frac{k_B T}{2} \sum_{|\ell\rangle} \sum_{|m\rangle} \tilde{a}^T_{m\ell} \left(\tilde{p}_\ell - \tilde{p}_m\right) \left(\ln \tilde{p}_\ell - \ln \tilde{p}_m\right) \tag{14.42}$$

leading to the spontaneous decrease of the Helmholtz free energy for a system in contact with a thermostat, a result we had from classical thermodynamics,

$$\frac{\mathrm{d}F}{\mathrm{d}t} \leqslant 0 \tag{14.43}$$

Here again, we can find the form of the equilibrium distribution by writing that at equilibrium, $F$ has reached a minimum, leading to $\tilde{p}_m = \tilde{p}_\ell$ for all pairs $(|\ell\rangle, |m\rangle)$. This yields

$$\frac{p_\ell}{p_m} = \frac{\tilde{p}_\ell \exp\left(-\dfrac{E_\ell}{k_B T}\right)}{\tilde{p}_m \exp\left(-\dfrac{E_m}{k_B T}\right)} = \exp\left(-\frac{E_\ell - E_m}{k_B T}\right) \tag{14.44}$$

which is the canonical distribution, as it should. Summarizing, for a system in contact with a thermostat that is initially out of equilibrium, the spontaneous evolution is in the sense of a decrease of the Helmholtz free energy, which reaches a minimum at equilibrium, at which point the probability distribution of the accessible states is canonical.

### 14.2.3 Irreversibility

Let us consider an isolated system, for which the master equation is (14.6), and assume that by some means the system has been prepared in a macroscopic state characterized by the initial probability distribution $\{p^0_\ell\}$. Leaving it to evolve spontaneously, if $\{p^0_\ell\}$ was not equal to the microcanonical distribution, then the $H$ theorem states that $\{p_\ell(t)\}$ will evolve in the sense of an increasing $S$. This means that there is no possibility for $\{p_\ell(t)\}$ to return spontaneously to $\{p^0_\ell\}$, ever. This irreversibility stems from the first-order in time master equation, and contrasts with the reversibility of microscopic processes, that are governed by **deterministic equations**. How one goes from these deterministic equations at the microscopic level to **stochastic (i.e., probabilistic) equations** at the macroscopic levels involves the extreme sensitivity to initial conditions and to very weak outside perturbations (e.g., gravitational or electromagnetic fields) : two trajectories in phase space that are extremely close initially diverge exponentially in time. A simple example of this is the gravitational influence of the observer on gas particle collisions occurring within a vessel. Assume that the vessel contains $1\,\mathrm{l}$ of nitrogen at normal conditions of temperature and pressure, and the observer of mass $M = 75\,\mathrm{kg}$ stands $D = 1\,\mathrm{m}$ away from the vessel. The gravitational interaction energy is extremely weak,

$$E_g = -\frac{GMM_{\mathrm{gas}}}{D^2} \sim \frac{2}{3}\,10^{-10}\frac{75}{1^2} \times \frac{28\,10^{-3}}{22.4} \sim -6\,10^{-12}\,\mathrm{J} \tag{14.45}$$

where we use the molar volume $22.4\mathrm{l} \cdot \mathrm{mol}^{-1}$ and the molar mass of nitrogen $28\,\mathrm{g} \cdot \mathrm{mol}^{-1}$. This energy is much lower than the internal energy

$$U = \frac{5}{2} N k_B T \sim \frac{5}{2} \times \frac{6.02\,10^{23}}{22.4} \times 1.38\,10^{-23} \times 300 \sim 250\,\mathrm{J} \tag{14.46}$$

but the tidal effect on two particles initially set for a collision perturbs that trajectory by an angle $\delta\theta_0$, which may be expressed as [8]

$$\delta\theta_0 \sim \frac{GM}{D^3} \left(\frac{\ell}{v}\right)^2 \tag{14.47}$$

8. See "Eléments de Physique Statistique" (B. Diu, C. Guthmann, D. Lederer, and B. Roulet), IV.B.

where $\ell$ is the mean free path ($\sim 140\,\mathrm{nm}$) and $v$ is the relative velocity between the two particles. This angle is very small, $\delta\theta_0 \sim 4\,10^{-28}\,\mathrm{rad}$, but in the collision, the change in direction is of order $\delta\theta_1 \sim \delta\theta_0(\ell/b)$ where $b$ is the interaction radius ($\sim 0.15\,\mathrm{nm}$), and after $q$ collisions, we have

$$\delta\theta_q \sim \delta\theta_0 \left(\frac{\ell}{b}\right)^q \tag{14.48}$$

which typically reaches $\sim 1\,\mathrm{rad}$ after only $q \sim 10$ collisions, which corresponds to a typical timescale $\sim 2.5\,10^{-9}\,\mathrm{s}$. In a very short time, the particles have lost any memory of where they came from, which explains why the microscopic time-reversibility is lost and the equations of evolution in statistical physics are not time-reversible.

# Part III

# Additional topics

# 15

# Kinetic theory

The **kinetic theory** is an approach to understanding the properties of gases as a collection of a large number of particles in random motion. It is historically older than what we have seen so far, as it can be traced back to Bernoulli's *Hydrodynamica* in 1738. In its simplest form, the theory assumes that these particles (atoms and molecules) have no internal structure, that they can be safely assumed to be non-relativistic, that quantum effects may be neglected, and that the interaction between particles is limited to binary elastic collisions. From this theory, we may derive the **equation of state** for ideal gases [1], and **transport coefficients** such as viscosity or thermal conductivity.

## 15.1 Lagrangian and Hamiltonian formalisms

### 15.1.1 Generalized coordinates and velocities

In section 1.5, we wrote how a macroscopic system may be described classically by the set of positions and momenta of its $N$ particles. More generally, the variables describing the state of the system will be the **generalized coordinates** $\vec{q} = \{q_i\}_{1 \leqslant i \leqslant s}$ and associated **generalized velocities** $\dot{\vec{q}} = \{\dot{q}_i\}_{1 \leqslant i \leqslant s}$. An example of a generalized coordinate is the angle made by a pendulum with respect to the vertical axis. We have written $s$ to stand for the number of degrees of freedom. For example, in a monoatomic gas with $N$ particles, we obviously have $s = 3N$.

### 15.1.2 Lagrangian formalism

#### Lagrangian and action

The **Lagrangian formalism** of classical mechanics amounts to writing the equations of motion in terms of the **Lagrange function** or **Lagrangian**, defined as the difference between the total kinetic energy $K$ and the potential energy $U$ from which the forces derive [2]

$$\mathcal{L}(\vec{q}, \dot{\vec{q}}, t) = K - U \tag{15.1}$$

Using this function, it is possible to define the **action** along a trajectory $\vec{q}(t)$ in **configuration space** [3],

1. Taking into account more realistic binary interactions, we may derive more general equations of state, in particular the Van der Waals equation.
2. This assumes that all the forces acting on the particles do derive from such an energy.
3. This is the $s$-dimensional space where $\vec{q}$ lives, not to be confused with the $2s$-dimensional **phase space**.

given a starting point $A$ at time $t_0$ and ending point $B$ reached at time $t_1$, as the following integral, which is a **functional** (i.e., a function whose argument is a function),

$$\mathcal{S}[\vec{q}(t)] = \int \mathcal{L}(\vec{q}, \dot{\vec{q}}, t)\mathrm{d}t \tag{15.2}$$

The search for the trajectory $\vec{q}_0(t)$ actually followed by the system is based on the **principle of stationary action**, also known as **principle of least action** or **Maupertuis' principle**, which states that for this trajectory, the action should be stationary, i.e., that a small variation $\delta\vec{q}(t)$ in the trajectory leads to a null variation of $\mathcal{S}$ to linear order, i.e.

$$\delta\mathcal{S} = \mathcal{S}[\vec{q}_0(t) + \delta\vec{q}(t)] - \mathcal{S}[\vec{q}_0(t)] = o\left[\delta\vec{q}(t)\right] \tag{15.3}$$

### Euler-Lagrange equations

This means that $\vec{q}_0(t)$ extremizes the action, and the calculus of variations[4] can be put to use to derive constraints on the Lagrangian, as we now show in the one-dimensional case. Assume $q_0(t)$ is a trajectory that extremizes the action and introduce a function $h(t)$ such that $h(t_0) = h(t_1) = 0$, where $t_0$ and $t_1$ are the initial and final times. With $\epsilon \ll 1$ a small parameter, we consider the alternate trajectories $q_\epsilon(t) = q_0(t) + \epsilon h(t)$ and write the action as a function of $\epsilon$

$$\mathcal{S}_\epsilon = \mathcal{S}\left[q_\epsilon(t)\right] = \int_{t_0}^{t_1} \mathcal{L}(q_\epsilon, \dot{q}_\epsilon, t)\mathrm{d}t \tag{15.4}$$

The stationary action principle suggests to write the derivative of this action with respect to $\epsilon$

$$\frac{\mathrm{d}\mathcal{S}_\epsilon}{\mathrm{d}\epsilon} = \int_{t_0}^{t_1} \frac{\mathrm{d}\mathcal{L}(q_\epsilon, \dot{q}_\epsilon, t)}{\mathrm{d}\epsilon}\mathrm{d}t = \int_{t_0}^{t_1} \left[\frac{\partial\mathcal{L}}{\partial q_\epsilon}\frac{\partial q_\epsilon}{\partial\epsilon} + \frac{\partial\mathcal{L}}{\partial\dot{q}_\epsilon}\frac{\partial\dot{q}_\epsilon}{\partial\epsilon}\right]\mathrm{d}t = \int_{t_0}^{t_1} \left[\frac{\partial\mathcal{L}}{\partial q_\epsilon}h + \frac{\partial\mathcal{L}}{\partial\dot{q}_\epsilon}\dot{h}\right]\mathrm{d}t \tag{15.5}$$

since we have

$$\frac{\partial q_\epsilon}{\partial\epsilon} = h \qquad\qquad \frac{\partial\dot{q}_\epsilon}{\partial\epsilon} = \dot{h} \tag{15.6}$$

This derivative should be zero for $\epsilon = 0$, which yields

$$\int_{t_0}^{t_1} \left[\frac{\partial\mathcal{L}}{\partial q_0}h + \frac{\partial\mathcal{L}}{\partial\dot{q}_0}\dot{h}\right]\mathrm{d}t = 0 \tag{15.7}$$

given that we have the following limits

$$\lim_{\epsilon\to 0}\frac{\partial\mathcal{L}}{\partial q_\epsilon} = \frac{\partial\mathcal{L}}{\partial q_0} \qquad\qquad \lim_{\epsilon\to 0}\frac{\partial\mathcal{L}}{\partial\dot{q}_\epsilon} = \frac{\partial\mathcal{L}}{\partial\dot{q}_0} \tag{15.8}$$

Performing an integration by parts on the second term, we get

$$\int_{t_0}^{t_1} h\left[\frac{\partial\mathcal{L}}{\partial q_0} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}_0}\right)\right]\mathrm{d}t + \left[h\frac{\partial\mathcal{L}}{\partial\dot{q}_0}\right]_{t_0}^{t_1} = 0 \tag{15.9}$$

but the fully-integrated term is null, due to the constraints $h(t_0) = h(t_1) = 0$, so we have

$$\int_{t_0}^{t_1} h\left[\frac{\partial\mathcal{L}}{\partial q_0} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}_0}\right)\right]\mathrm{d}t = 0 \tag{15.10}$$

4. The whole field can be traced back to Euler and Lagrange searching for the solution of the **tautochrone curve** problem, and has wide extensions to optics (**Fermat's principle**) and quantum field theory (**Feynman's path integral formulation**).

and this must be true whatever the "shape" of the disturbance curve $h(t)$. Consequently [5], the expression inside the square brackets must be identically zero. Generalizing to arbitrary dimensions, we obtain the equations of motion in the form of the **Euler-Lagrange equations**

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) = \frac{\partial \mathcal{L}}{\partial q_i} \tag{15.11}$$

#### Examples

For instance, take the case of a one-dimensional free harmonic oscillator, for which the kinetic energy is quadratic in $\dot{q}$ and the potential energy is quadratic in $q$,

$$\mathcal{L}(q,\dot{q}) = K(\dot{q}) - U(q) = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}kq^2 \tag{15.12}$$

Then we have

$$\frac{\partial \mathcal{L}}{\partial \dot{q}} = \frac{\partial K}{\partial \dot{q}} = m\dot{q} \qquad \frac{\partial \mathcal{L}}{\partial q} = -\frac{\partial U}{\partial q} = -kq \tag{15.13}$$

and the Lagrange equation correctly recovers the usual equation of motion

$$m\ddot{q} + kq = 0 \tag{15.14}$$

Another example of a Lagrangian is that governing the evolution of a system of charges in an electromagnetic field, which reads

$$\mathcal{L} = \sum_\alpha \left[K_\alpha - q_\alpha V_\alpha + q_\alpha \vec{A}_\alpha . \vec{v}_\alpha\right] \tag{15.15}$$

where the sum extends over particles indexed by $\alpha$. The kinetic energy of each particle is $K_\alpha$, while $V_\alpha$ and $\vec{A}_\alpha$ are the scalar and vector potentials, respectively, that apply at the position of the particle [6].

### 15.1.3 Hamiltonian formalism

#### Hamiltonian and Hamilton-Jacobi equations

The **Hamiltonian formalism** of classical mechanics is built upon the Lagrangian one by defining the **conjugate (generalized) momenta** of the generalized coordinates

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \tag{15.17}$$

and the **Hamilton function**, also known as **energy function** or **Hamiltonian**, that depends on the generalized coordinates and momenta

$$\mathcal{H}(\vec{q},\vec{p},t) = \sum_{i=1}^{s} p_i \dot{q}_i - \mathcal{L}(\vec{q},\dot{\vec{q}},t) \tag{15.18}$$

5. This is the fundamental lemma of the calculus of variations
6. We recall that these potentials are related to the electromagnetic field by

$$\vec{B} = \vec{\nabla} \times \vec{A} \qquad \vec{E} = -\vec{\nabla}V - \frac{\partial \vec{A}}{\partial t} \tag{15.16}$$

This is a **Legendre transform** from a function that depends on $(\vec{q}, \dot{\vec{q}})$ to one that depends on $(\vec{q}, \vec{p})$. In this formalism, the state of the system is given by the values of the $s$ generalized coordinates and $s$ generalized momenta, living in **phase space**. The equations of motion are found by taking the partial derivatives of $\mathcal{H}$ and making use of the Euler-Lagrange equations. By definition,

$$\frac{\partial \mathcal{H}}{\partial p_i} = \dot{q}_i \qquad \frac{\partial \mathcal{H}}{\partial q_i} = -\frac{\partial \mathcal{L}}{\partial q_i} = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) = -\dot{p}_i \tag{15.19}$$

The set of $s$ second-order Euler-Lagrange equations is therefore equivalent to the set of $2s$ first-order **Hamilton-Jacobi equations**

$$\frac{\mathrm{d}q_i}{\mathrm{d}t} = \frac{\partial \mathcal{H}}{\partial p_i} \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial q_i} \tag{15.20}$$

When forces derive from a potential energy $U(\vec{q})$ and when the kinetic energy is quadratic in all $\dot{q}_i$, the Hamiltonian identifies with the total energy. Indeed, we have

$$\mathcal{H} = \sum_{i=1}^{s} p_i \dot{q}_i - \mathcal{L} = \sum_{i=1}^{s} \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - K + U = \sum_{i=1}^{s} \dot{q}_i \frac{\partial K}{\partial \dot{q}_i} - K + U \tag{15.21}$$

and with $K$ quadratic in all $\dot{q}_i$, i.e.,

$$K = \sum_{i=1}^{s} a_i \dot{q}_i^2 \tag{15.22}$$

we have

$$\frac{\partial K}{\partial \dot{q}_i} = 2a_i \dot{q}_i \qquad \text{so} \qquad \sum_{i=1}^{s} \dot{q}_i \frac{\partial K}{\partial \dot{q}_i} = \sum_{i=1}^{s} 2a_i \dot{q}_i^2 = 2K \tag{15.23}$$

leading to the identity

$$\mathcal{H} = K + U \tag{15.24}$$

### Examples

As an example, take again the one-dimensional harmonic oscillator, for which

$$p = \frac{\partial \mathcal{L}}{\partial \dot{q}} = m\dot{q} \tag{15.25}$$

so that the Hamiltonian $\mathcal{H} = p\dot{q} - \mathcal{L}$ is

$$\mathcal{H} = m\dot{q}^2 - \left(\frac{1}{2}m\dot{q}^2 - \frac{1}{2}kq^2\right) = \frac{1}{2}m\dot{q}^2 + \frac{1}{2}kq^2 = \frac{p^2}{2m} + \frac{1}{2}kq^2 \tag{15.26}$$

The Hamilton-Jacobi equations then read, as they should,

$$\dot{q} = \frac{\partial \mathcal{H}}{\partial p} = \frac{p}{m} \qquad \dot{p} = -\frac{\partial \mathcal{H}}{\partial q} = -kq \tag{15.27}$$

For a system of charged particles in an electromagnetic field, the generalized momenta are [7]

$$\vec{p}_\alpha = \frac{\partial \mathcal{L}}{\partial \vec{v}_\alpha} = \frac{\partial K_\alpha}{\partial \vec{v}_\alpha} + q_\alpha \vec{A}_\alpha = m_\alpha \vec{v}_\alpha + q_\alpha \vec{A}_\alpha \tag{15.28}$$

7. The notation $\partial a / \partial \vec{b}$ stands for the vector whose components are $(\partial a/\partial b_x, \partial a/\partial b_y, \partial a/\partial b_z)$.

and the Hamiltonian is then

$$\mathcal{H} = \sum_\alpha \vec{p}_\alpha \cdot \vec{v}_\alpha - \mathcal{L} = \sum_\alpha \left( m_\alpha v_\alpha^2 + q_\alpha \vec{A}_\alpha \cdot \vec{v}_\alpha \right) - \left\{ \sum_\alpha \left[ K_\alpha - q_\alpha V_\alpha + q_\alpha \vec{A}_\alpha . \vec{v}_\alpha \right] \right\} \tag{15.29}$$

In the first sum we have twice the kinetic energy, and the terms in $q_\alpha \vec{A}_\alpha . \vec{v}_\alpha$ cancel out, leading to

$$\mathcal{H} = \sum_\alpha \left( \frac{1}{2} m_\alpha v_\alpha^2 + q_\alpha V_\alpha \right) \tag{15.30}$$

which is more usually written in terms of the momenta as

$$\mathcal{H} = \sum_\alpha \left[ \frac{\left( \vec{p}_\alpha - q_\alpha \vec{A}_\alpha \right)^2}{2m_\alpha} + q_\alpha V_\alpha \right] \tag{15.31}$$

### 15.1.4 Liouville's theorem

#### Conservation of the volume of phase space

The state of the system being given by the values of the $s$ generalized coordinates and momenta, it may be represented as a point in the $2s$-dimensional phase space, as shown in Fig. 15.1. This point will move in time due to the motions and interactions of the particles among themselves, and so we may speak of the "trajectory" of the system in phase space. Let us consider an elementary volume $\mathrm{d}\tau = \mathrm{d}\vec{q}\mathrm{d}\vec{p}$ in this phase space, and a set of systems whose representative points fill that elementary volume at a given instant. The equations of motion specify how these points will move in phase space over a certain short interval $\delta t$, to end up filling another elementary volume $\mathrm{d}\tau' = \mathrm{d}\vec{q}'\mathrm{d}\vec{p}'$. It may be shown [8] that to first order in $\delta t$, the elementary volume in phase space is conserved, a result known as **Liouville's theorem**,

$$\mathrm{d}\vec{q}'\mathrm{d}\vec{p}' = \mathrm{d}\vec{q}\mathrm{d}\vec{p} \tag{15.32}$$

This result generalizes to arbitrary times by integration.

#### Invariance of the distribution function

Liouville's theorem has an important consequence for the classical statistical description of macroscopic systems. We recall that the probability of finding the system in a state $(\vec{q}, \vec{p})$ at time $t$ is written

$$\mathrm{d}P(\vec{q}, \vec{p}, t) = f_N(\vec{q}, \vec{p}, t)\mathrm{d}\vec{q}\mathrm{d}\vec{p} \tag{15.33}$$

This distribution function $f_N$ is called $s$-**dimensional** because it depends on all $s$ coordinates and momenta. It may also be called $N$-**particle distribution function** to emphasize the number of particles $N$. We recall that $s = 3N$ here. Now, consider $\mathcal{Q}$ identical systems. The number of such systems whose representative points lie in the elementary volume $\mathrm{d}\vec{q}\mathrm{d}\vec{p}$ at time $t$ is, by definition, $\mathcal{Q} f_N(\vec{q}, \vec{p}, t)\mathrm{d}\vec{q}\mathrm{d}\vec{p}$. After $\delta t$, these systems will lie in the elementary volume $\mathrm{d}\vec{q'}\mathrm{d}\vec{p'}$, and by definition, their number is also $\mathcal{Q} f_N(\vec{q'}, \vec{p'}, t + \delta t)\mathrm{d}\vec{q'}\mathrm{d}\vec{p'}$. Since these are the same systems, and given Liouville's theorem, we have

$$f_N(\vec{q'}, \vec{p'}, t + \delta t) = f_N(\vec{q}, \vec{p}, t) \tag{15.34}$$

8. An elegant demonstration of this result amounts to showing the "incompressibility of the probability density fluid", and that this condition, written as $\vec{\nabla} \cdot \vec{v} = 0$ stems from the Hamilton-Jacobi equations.

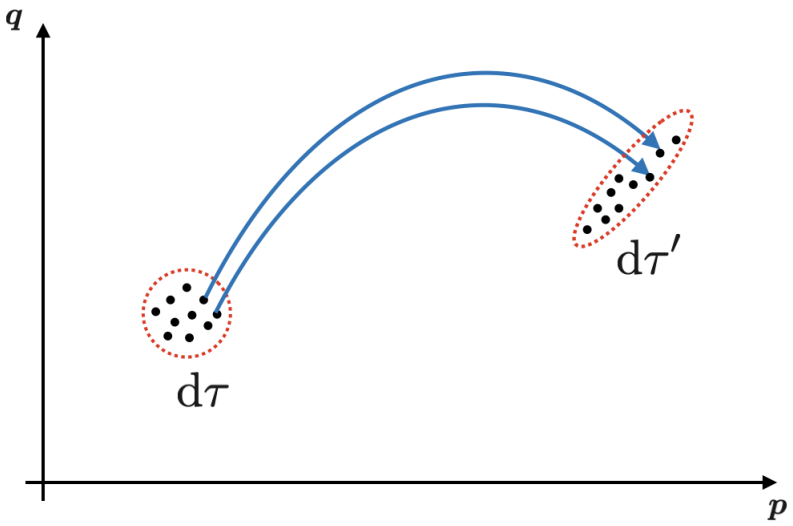



FIGURE 15.1 – Time evolution of an element of phase space, whose volume is conserved, $\mathrm{d}\tau' = \mathrm{d}\tau$. Each point represents a system, evolving along the blue trajectories. Note that these cannot intersect.

which states that the $s$-dimensional distribution function is a constant along trajectories in phase-space. We can show that this is also a relativistic invariant, a result that is useful to discuss the properties of the spectrum of the cosmic microwave background (CMB).

### Liouville's equation

Writing this conservation of $f_N\left[\vec{q}(t), \vec{p}(t), t\right]$ explicitly, we obtain **Liouville's equation**

$$\frac{\mathrm{d}f_N}{\mathrm{d}t} = \frac{\partial f_N}{\partial t} + \sum_{i=1}^{s}\left(\frac{\partial f_N}{\partial q_i}\dot{q}_i + \frac{\partial f_N}{\partial p_i}\dot{p}_i\right) = 0 \tag{15.35}$$

This may be viewed as the conservation of a density in a flow of particles in a $2s$-dimensional space. Indeed, in this space, the "velocity" $\vec{v}$ of the system has coordinates $(\dot{q}_1, \ldots, \dot{q}_s, \dot{p}_1, \ldots, \dot{p}_s)$ and the partial derivatives of $f_N$ with respect to the coordinates in phase space form a gradient of $f_N$, so

$$\frac{\mathrm{d}f_N}{\mathrm{d}t} = \frac{\partial f_N}{\partial t} + \vec{v}\cdot\vec{\nabla} f_N = 0 \tag{15.36}$$

which is just like a conservation equation for an advected quantity $f_N$ using a Lagrangian approach in hydrodynamics (i.e., following the mesoscopic fluid particles).

### Poisson brackets

If the system is macroscopically at equilibrium, this density must be stationary, i.e., it cannot depend explicitly on time, so we have the situation described in Fig. 15.2. Mathematically, this means that

$$\frac{\partial f_N}{\partial t} = 0 \qquad \sum_{i=1}^{s}\left(\frac{\partial f_N}{\partial q_i}\dot{q}_i + \frac{\partial f_N}{\partial p_i}\dot{p}_i\right) = 0 \tag{15.37}$$

Note that the Hamilton-Jacobi equations allow Liouville's equation to be rewritten as

$$\frac{\mathrm{d}f_N}{\mathrm{d}t} = \frac{\partial f_N}{\partial t} + \{f_N, \mathcal{H}\} = 0 \tag{15.38}$$

introducing the **Poisson bracket** defined by

$$\{A, B\} = \sum_{i=1}^{s} \left[ \frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} \right] \tag{15.39}$$

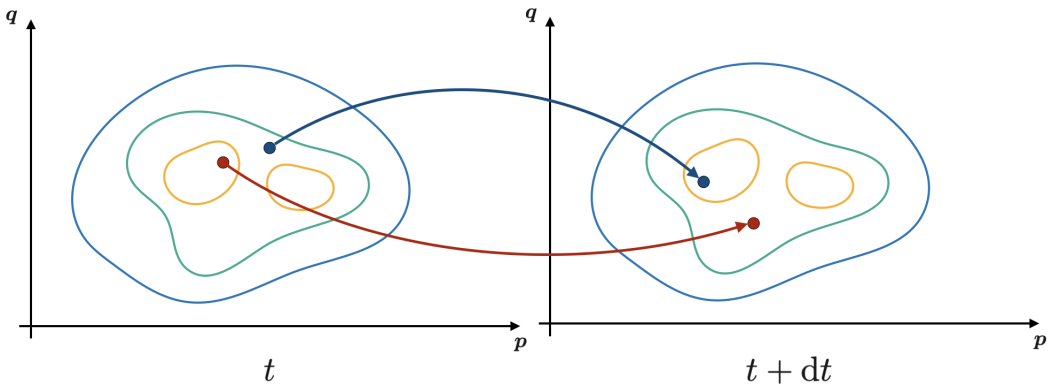



FIGURE 15.2 – Time evolution of the density $f_N$ for a system that is macroscopically at rest, between $t$ and $t + \mathrm{d}t$. Coloured contours represent isodensity curves, and two representative points are shown. They move in phase space but the overall density of points does not change.

For a system macroscopically at rest, where the $s$-dimensional distribution function does not explicitly depend on time, we then have

$$\{f_N, \mathcal{H}\} = 0 \tag{15.40}$$

In fact, for any observable $A$ that depends on the generalized coordinates and momenta and on time, we have the total time derivative as

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{\partial A}{\partial t} + \{A, \mathcal{H}\} \tag{15.41}$$

from the same development that led to Eq. (15.35). For example, the time evolution of the energy (i.e., the Hamiltonian) is just its explicit time dependence

$$\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t} = \frac{\partial \mathcal{H}}{\partial t} \tag{15.42}$$

Note that the Poisson brackets have a quantum-mechanical counterpart that is the **commutator** between two operators

$$[A, B] = AB - BA \tag{15.43}$$

and the time evolution equation above has an equivalent in the **Ehrenfest theorem** that expresses the evolution of the expectation value of any quantum mechanical operator $A$

$$\frac{\mathrm{d}\langle A \rangle}{\mathrm{d}t} = \left\langle \frac{\partial A}{\partial t} \right\rangle + \frac{1}{i\hbar} \langle [A, H] \rangle \tag{15.44}$$

where $H$ is the (quantum) Hamiltonian of the system. An application of that theorem are the **Ehrenfest relations** stating that the average values of the position and momentum operators obey the classical Newton relations,

$$\frac{\mathrm{d}\langle\vec{X}\rangle}{\mathrm{d}t} = \frac{\langle\vec{P}\rangle}{m} \qquad \frac{\mathrm{d}\langle\vec{P}\rangle}{\mathrm{d}t} = \langle\vec{F}\rangle \tag{15.45}$$

## 15.2 Distribution functions

### 15.2.1 BBGKY hierarchy

For a set of $N$ point-like particles, we can take their positions $\vec{r}_i$ to be the generalized coordinates $\vec{q}_i$, in which case the generalized momenta are simply $\vec{p}_i = m\vec{v}_i$, where $m$ is the mass of a particle. Then, Liouville's equation may be written in a form that combines terms pertaining to the same particle $i$ and explicitly makes use of the equations of motion, as

$$\frac{\partial f_N}{\partial t} + \sum_{i=1}^{N}\left(\frac{\vec{p}_i}{m}.\frac{\partial f_N}{\partial \vec{q}_i} + \vec{F}_i.\frac{\partial f_N}{\partial \vec{p}_i}\right) = 0 \tag{15.46}$$

In this equation, the net force exerted on particle $i$ is

$$\vec{F}_i = -\frac{\partial \mathcal{U}_i}{\partial \vec{q}_i} = -\sum_{j\neq i}\frac{\partial U_{ij}}{\partial \vec{q}_i} - \frac{\partial U_{\mathrm{ext}}}{\partial \vec{q}_i} \tag{15.47}$$

where $U_{ij}$ is the interaction potential between particles $i$ and $j$, $U_{\mathrm{ext}}$ is the externally-imposed potential, and the total potential exerted on particle $i$ is

$$\mathcal{U}_i = U_{\mathrm{ext}} + \sum_{j\neq i} U_{ij} \tag{15.48}$$

Solving Liouville's equation governing the evolution of $f_N$ in $6N$-dimensional phase space is impossible in the general case[9], but in fact $f_N$ contains much more information than what is necessary for a description of equilibrium properties, and it usually suffices to focus on the **single-particle distribution function** $f_1(\vec{q}_1, \vec{p}_1, t)$ living in 6-dimensional space. This function is defined by writing that the probability $\mathrm{d}P(\vec{q}_1, \vec{p}_1, t)$ to find any particle within the elementary volume $\mathrm{d}\vec{q}_1\mathrm{d}\vec{p}_1$ is

$$\mathrm{d}P(\vec{q}_1, \vec{p}_1, t) = f_1(\vec{q}_1, \vec{p}_1, t)\mathrm{d}\vec{q}_1\mathrm{d}\vec{p}_1 \tag{15.49}$$

It is straightforward to write $f_1$ as an integral of $f_N$ over the $N-1$ other particles, with

$$f_1(\vec{q}_1, \vec{p}_1, t) = \int f_N\left(\vec{q}_1, \vec{p}_1, \vec{q}_2, \vec{p}_2, \ldots, t\right)\prod_{i=2}^{N}\mathrm{d}\vec{q}_i\mathrm{d}\vec{p}_i \tag{15.50}$$

This process of integrating over a subset of particles is called **marginalization**, a term originating in probability theory. Similarly, a two-particle distribution function $f_2$ may be defined as

$$f_2(\vec{q}_1, \vec{p}_1, \vec{q}_2, \vec{p}_2, t) = \int f_N\left(\vec{q}_1, \vec{p}_1, \vec{q}_2, \vec{p}_2, \ldots, t\right)\prod_{i=3}^{N}\mathrm{d}\vec{q}_i\mathrm{d}\vec{p}_i \tag{15.51}$$

9. This would amount basically to knowing the full details of the microscopic state of the system, an endeavour we have renounced from the very beginning.

and, more generally, an $r$-particle distribution function $f_r$ (with $r < N$) would be defined by

$$f_r\left(\vec{q}_1, \ldots, \vec{p}_r, t\right) = \int f_N\left(\vec{q}_1, \vec{p}_1, \vec{q}_2, \vec{p}_2, \ldots, t\right) \prod_{i=r+1}^{N} \mathrm{d}\vec{q}_i \mathrm{d}\vec{p}_i \tag{15.52}$$

It may be shown [10] that the time evolution of the $r$-particle distribution function $f_r$ reads [11]

$$\frac{\partial f_r}{\partial t} + \{f_r, \mathcal{H}_r\} = (N-r) \sum_{i=1}^{r} \int \left( \frac{\partial U_{ir+1}}{\partial \vec{q}_i} \cdot \frac{\partial f_{r+1}}{\partial \vec{p}_i} \right) \mathrm{d}\vec{q}_{r+1} \mathrm{d}\vec{p}_{r+1} \tag{15.53}$$

where the Hamiltonian $\mathcal{H}_r$ only includes interactions involving the first $r$ particles, i.e.,

$$\mathcal{H}_r = \sum_{i=1}^{r} \left[ \frac{p_i^2}{2m} + U_{\mathrm{ext}}(\vec{q}_i) \right] + \frac{1}{2} \sum_{i=1}^{r} \sum_{j=1}^{r} U_{ij}(\vec{q}_i, \vec{q}_j) \tag{15.54}$$

With this subset Hamiltonian, the Poisson bracket reads

$$\{f_r, \mathcal{H}_r\} = \sum_{i=1}^{r} \frac{\vec{p}_i}{m} \cdot \frac{\partial f_r}{\partial \vec{q}_i} - \sum_{i=1}^{r} \left[ \left( \sum_{\substack{j=1 \\ j \neq i}}^{r} \frac{\partial U_{ij}}{\partial \vec{q}_i} \right) + \frac{\partial U_{\mathrm{ext}}}{\partial \vec{q}_i} \right] \cdot \frac{\partial f_r}{\partial \vec{p}_i} \tag{15.55}$$

Equation (15.53) shows that the evolution of the $r$-particle distribution function depends on the properties of the $(r+1)$-particle distribution function $f_{r+1}$, that appears under the integral on the right-hand side. For instance, the first of these equations, for $r = 1$, reads

$$\frac{\partial f_1}{\partial t} + \frac{\vec{p}_1}{m} \cdot \frac{\partial f_1}{\partial \vec{q}_1} - \frac{\partial U_{\mathrm{ext}}}{\partial \vec{q}_1} \cdot \frac{\partial f_1}{\partial \vec{p}_1} = (N-1) \int \frac{\partial U_{12}}{\partial \vec{q}_1} \frac{\partial f_2}{\partial \vec{p}_1} \mathrm{d}\vec{q}_2 \mathrm{d}\vec{p}_2 \tag{15.56}$$

whose right-hand side involves the two-particle distribution function $f_2(\vec{q}_1, \vec{p}_1, \vec{q}_2, \vec{p}_2, t)$. To solve the above equation requires knowing this new function, which obeys a similar equation but where the right-hand side now involves the 3-particle distribution function $f_3$, and so on. This set of equations constitutes the **BBGKY hierarchy**, named after Bogoliubov, Born, Green, Kirkwood, and Yvon.

### 15.2.2 Boltzmann equation

The BBGKY hierarchy is a moment approach, and as is usual in such cases, a simplification of the problem requires truncating the hierarchy of moments. This means that if a model for $f_{r+1}$ is available, one can use it to solve for $f_r$ and go back down the hierarchy to solve for $f_1$. The **Boltzmann equation** is one such approach, that models $f_2$ based on the hypothesis of **molecular chaos**, also known as **Stosszahlansatz**, which is the assumption that the velocities of colliding particles are uncorrelated, and do not depend on their positions. In that case, the equation modelling the evolution of $f_2$ is truncated by setting the corresponding right-hand side (involving $f_3$) to zero [12]. It is also reasonable to expect that at large distances [13], the probability density $f_2$ should factorize [14]

$$f_2(\vec{q}_1, \vec{p}_1, \vec{q}_2, \vec{p}_2, t) \approx f_1(\vec{q}_1, \vec{p}_1, t) f_1(\vec{q}_2, \vec{p}_2, t) \tag{15.57}$$

10. See for instance M. Kardar's course, available at `http://web.mit.edu/8.333/www/lectures/`, section III.C
11. Note that for $r = N$ we recover the conservation of $f_N$.
12. This is equivalent to ignoring collisions with more than two partners.
13. By which we mean that $|\vec{q}_2 - \vec{q}_1| \gg d$, where $d$ is the characteristic length scale for the interparticular interaction.
14. Recall that this is what happens for probability distributions of *independent* random variables.

A more detailed discussion of the right-hand side terms of Eq. (15.56) yields that one may write

$$\frac{\partial f_1}{\partial t} + \frac{\vec{p}_1}{m} \cdot \frac{\partial f_1}{\partial \vec{q}_1} - \frac{\partial U_{\text{ext}}}{\partial \vec{q}_1} \cdot \frac{\partial f_1}{\partial \vec{p}_1} = \left( \frac{D f_1}{D t} \right)_{\text{coll}} \tag{15.58}$$

which is the **Boltzmann equation**, and where the right-hand side is the **collisional integral** characterizing the redistribution of particles in phase space due to their mutual interaction through collisions,

$$-\left( \frac{D f_1}{D t} \right)_{\text{coll}} = \iint \left[ f_1\left(\vec{q}_1, \vec{p}_1\right) f_1\left(\vec{q}_1, \vec{p}_2\right) - f_1\left(\vec{q}_1, \vec{p'}_1\right) f_1\left(\vec{q}_1, \vec{p'}_2\right) \right] |\vec{v}_2 - \vec{v}_1| \left| \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \right| \mathrm{d}\Omega \mathrm{d}^3 \vec{p}_2 \tag{15.59}$$

For simplicity, the dependence of $f_1$ on time was omitted. The quantity $\mathrm{d}\sigma/\mathrm{d}\Omega$ is the **differential cross-section** characterizing the probability for a collision (see Fig. 15.3) in which partners move from a state $(\vec{p}_1, \vec{p}_2)$ to $\left(\vec{p'}_1, \vec{p'}_2\right)$. Overall, the molecular chaos hypothesis assumes that the 2-particle distribution function may be expressed in terms of products of the single-particle one. The meaning of this expression for the collisional integral is that the probability to find a particle at $\vec{q}_1$ with momentum $\vec{p}_1$ is suddenly altered by a collision with a partner having momentum $\vec{p}_2$. The probability of such a collision involves the differential cross-section, the relative velocity $|\vec{v}_2 - \vec{v}_1|$, and the joint probability of finding two particles with these momenta at that position. The first term in the integral represents a probability subtraction due to particles being ejected from the phase space point $(\vec{q}_1, \vec{p}_1)$ in the collision. The second term represents the inverse process [15].

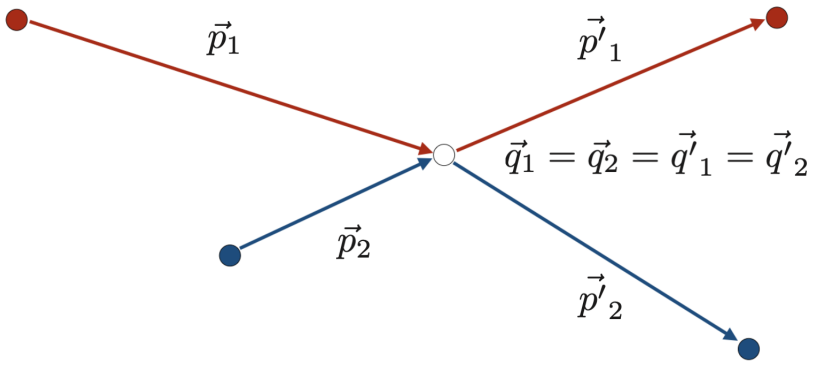



FIGURE 15.3 – Schematic two-body collision occurring at a point in space $\vec{q}_1 = \vec{q}_2 = \vec{q'}_1 = \vec{q'}_2$, where partners with ingoing momenta $\vec{p}_1$ and $\vec{p}_2$ leave with momenta $\vec{p'}_1$ and $\vec{p'}_2$. These events are the ones taken into account in the collisional integral (15.59).

### 15.2.3 Vlasov equations

In the case of interactions among particles that do not reduce to collisions, such as long-range Coulomb forces in a plasma, or gravitational forces in a stellar system (globular clusters or galaxies), difficulties appear with the Boltzmann equation as written in Eq. (15.58). In that situation, Vlasov proposed to consider the system as **collisionless** and to include the interactions as resulting from a **mean field** for each particle. For example, in the case of a fully ionized plasma of electrons and ions with charge $Ze$, their respective distribution functions $f_e(\vec{q}, \vec{p}, t)$ and $f_i(\vec{q}, \vec{p}, t)$ obey the following system of equations

15. Note that $\vec{p'}_1$ and $\vec{p'}_2$ are both functions of $\vec{p}_1$ and $\vec{p}_2$.

$$\frac{\partial f_e}{\partial t} + \vec{v}_e \cdot \frac{\partial f_e}{\partial \vec{q}} - e\left(\vec{E} + \vec{v}_e \times \vec{B}\right) \cdot \frac{\partial f_e}{\partial \vec{p}} = 0 \tag{15.60}$$

$$\frac{\partial f_i}{\partial t} + \vec{v}_i \cdot \frac{\partial f_i}{\partial \vec{q}} + Ze\left(\vec{E} + \vec{v}_i \times \vec{B}\right) \cdot \frac{\partial f_i}{\partial \vec{p}} = 0 \tag{15.61}$$

that is to be completed by the specification of the electromagnetic field that arises from the distributions of electrons and ions. This is done by way of the Maxwell equations

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0} \qquad \vec{\nabla} \cdot \vec{B} = 0 \qquad \vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \qquad \vec{\nabla} \times \vec{B} = \mu_0 \vec{j} + \frac{1}{c^2}\frac{\partial \vec{E}}{\partial t} \tag{15.62}$$

where the distribution functions are hidden in the charge and current densities, since

$$\rho = e \int (Z f_i - f_e)\, \mathrm{d}^3\vec{p} \qquad \vec{j} = e \int (Z f_i \vec{v}_i - f_e \vec{v}_e)\, \mathrm{d}^3\vec{p} \tag{15.63}$$

This whole set of equations, Eqs. (15.60) to (15.63), constitute the **Vlasov-Maxwell** system. In the non-relativistic zero-magnetic field limit, this becomes the **Vlasov-Poisson equations**

$$\frac{\partial f_e}{\partial t} + \vec{v}_e \cdot \frac{\partial f_e}{\partial \vec{q}} - e\vec{E} \cdot \frac{\partial f_e}{\partial \vec{p}} = 0 \qquad \frac{\partial f_i}{\partial t} + \vec{v}_i \cdot \frac{\partial f_i}{\partial \vec{q}} + Ze\vec{E} \cdot \frac{\partial f_i}{\partial \vec{p}} = 0 \tag{15.64}$$

$$\vec{\nabla} \cdot \vec{E} = \frac{\rho}{\epsilon_0} \qquad \rho = e \int (Z f_i - f_e)\, \mathrm{d}^3\vec{p} \tag{15.65}$$

### 15.2.4 Maxwell distribution

#### Equilibrium distribution of the Boltzmann equation

It may be shown that the equilibrium occurs if and only if the collisional integral in the Boltzmann equation is null. It follows that the one-particle distribution function must be such that

$$f_1^{(0)}(\vec{p}_1) f_1^{(0)}(\vec{p}_2) = f_1^{(0)}(\vec{p'}_1) f_1^{(0)}(\vec{p'}_2) \tag{15.66}$$

omitting the dependency on $\vec{q}_1$ since it appears identically in all these functions. We have used a "(0)" superscript to mark that these one-particle distribution functions are those at equilibrium. By taking the logarithm, we have

$$\ln f_1^{(0)}(\vec{p'}_1) + \ln f_1^{(0)}(\vec{p'}_2) = \ln f_1^{(0)}(\vec{p}_1) + \ln f_1^{(0)}(\vec{p}_2) \tag{15.67}$$

which shows that $\ln f_1^{(0)}$ is a **collisional invariant**, which means it is a conserved quantity in a two-body collision. It should then be a linear combination of the only collisional invariants there is, i.e., the total mass, momentum, and kinetic energy.

#### Determination of the distribution from general considerations

We have not developed too much the above reasoning, nor given the form of the equilibrium distribution $f_1^{(0)}$, to focus on a demonstration that makes use of very general considerations (stationarity, homogeneity, and isotropy) for a macroscopic system in thermal equilibrium at temperature $T$. We start from the probability $\mathrm{d}P(\vec{r}, \vec{p}, t)$ to find, at time $t$, a particle within the element $\mathrm{d}^3\vec{r}\mathrm{d}^3\vec{p}$ around $(\vec{r}, \vec{p})$ in phase space, i.e.,

$$\mathrm{d}P(\vec{r}, \vec{p}, t) = f_1(\vec{r}, \vec{p}, t)\mathrm{d}^3\vec{r}\mathrm{d}^3\vec{p} \tag{15.68}$$

The stationarity hypothesis means that $f_1$ cannot explicitly depend on time $t$, and the homogeneity means it does not depend on position $\vec{r}$ either, so

$$\mathrm{d}P(\vec{r},\vec{p},t) = f_1(\vec{p})\mathrm{d}^3\vec{r}\mathrm{d}^3\vec{p}. \tag{15.69}$$

By integrating over position, we get a distribution function for the sole momentum $f(\vec{p})$

$$f(\vec{p}) = \int f_1(\vec{p})\mathrm{d}^3\vec{r} = V f_1(\vec{p}), \tag{15.70}$$

where $V$ is the volume of the system. The isotropy hypothesis implies that the distribution can only depend on the modulus $p$ of $\vec{p}$, or equivalently on $p^2$, so there is a function $g$ such that

$$f(\vec{p}) = g(p^2) = g(p_x^2+p_y^2+p_z^2) = h(p_x^2)h(p_y^2)h(p_z^2). \tag{15.71}$$

the last step stemming from the fact that all three components of the momentum are independent random variables that are identically distributed. The independence implies that the probability distribution $g$ may be factorized, the identical laws of probability implying that the factors are the same function $h$. Taking the logarithm of the above relation and deriving with respect to one component, say $p_x$, we get

$$\frac{\partial \ln g}{\partial p_x} = \frac{g'(p^2)}{g(p^2)}\frac{\partial p^2}{\partial p_x} = 2p_x\frac{g'(p^2)}{g(p^2)} = \frac{\partial \ln h(p_x^2)}{\partial p_x} = 2p_x\frac{h'(p_x^2)}{h(p_x^2)}. \tag{15.72}$$

from which we have, using similar derivations with respect to $p_y$ and $p_z$,

$$\frac{g'(p^2)}{g(p^2)} = \frac{h'(p_x^2)}{h(p_x^2)} = \frac{h'(p_y^2)}{h(p_y^2)} = \frac{h'(p_z^2)}{h(p_z^2)} = -B, \tag{15.73}$$

where $B$ is a constant [16]. Integrating, we get

$$h(p_x^2) = A\exp\left(-Bp_x^2\right) \qquad g(p^2) = A^3\exp\left(-Bp^2\right), \tag{15.74}$$

which shows that the distribution function for the momenta, and from there the velocities, is a **Gaussian**. The $A$ and $B$ constants are determined from the normalization of probabilities and from the relation between the mean kinetic energy and the temperature (see 3.5.2)

$$\frac{\langle p^2\rangle}{2m} = \frac{3k_BT}{2} \tag{15.75}$$

Indeed, we have [17]

$$\int f(\vec{p})\mathrm{d}^3\vec{p} = \int_0^\infty 4\pi p^2 g(p^2)\mathrm{d}p = 4\pi A^3\int_0^\infty p^2\exp\left(-Bp^2\right)\mathrm{d}p = 4\pi A^3\frac{\sqrt{\pi}}{4B^{3/2}} = 1 \tag{15.77}$$

which gives $A = \sqrt{B/\pi}$. The mean kinetic energy is given by the integral

$$\frac{\langle p^2\rangle}{2m} = \frac{1}{2m}\int p^2 f(\vec{p})\mathrm{d}^3\vec{p} = \frac{4\pi A^3}{2m}\int_0^\infty p^4\exp\left(-Bp^2\right)\mathrm{d}p = \frac{4\pi A^3}{2m}\frac{3}{8B^2}\sqrt{\frac{\pi}{B}} = \frac{3k_BT}{2} \tag{15.78}$$

16. For any values of $p_x$ and $p_y$, and writing $F(u) = h'(u^2)/h(u^2)$ we have $F(p_x) = F(p_y)$ which is precisely what a constant function is.

17. We note the following useful integrals

$$\int_0^\infty x^{2n}e^{-bx^2}\mathrm{d}x = \frac{(2n-1)!!}{b^n 2^{n+1}}\sqrt{\frac{\pi}{b}} \qquad \int_0^\infty x^{2n+1}e^{-bx^2}\mathrm{d}x = \frac{n!}{2b^{n+1}} \tag{15.76}$$

Using the relation between $A$ and $B$, this yields

$$\frac{4\pi B^{3/2}}{2m\pi^{3/2}}\frac{3}{8B^2}\sqrt{\frac{\pi}{B}} = \frac{3}{4mB} = \frac{3k_BT}{2} \tag{15.79}$$

and so $B = 1/(2mk_BT)$ and we have the single-particle distribution function that takes the form of the **Maxwell distribution** or **Maxwell-Boltzmann distribution**

$$f(\vec{p}) = \left(\frac{1}{2\pi mk_BT}\right)^{3/2} \exp\left(-\frac{p^2}{2mk_BT}\right) \tag{15.80}$$

which can readily be written in terms of the velocities rather than the momenta

$$f(\vec{v}) = \left(\frac{m}{2\pi k_BT}\right)^{3/2} \exp\left(-\frac{mv^2}{2k_BT}\right). \tag{15.81}$$

such that the probability for a particle to have a velocity $\vec{v}$ to within $\mathrm{d}^3\vec{v}$ is $\mathrm{d}P(\vec{v}) = f(\vec{v})\mathrm{d}^3\vec{v}$. In equation (15.81), $v = ||\vec{v}||$ is the magnitude of the velocity vector. This distribution is therefore a 3-dimensional Gaussian probability for the vector $(v_x, v_y, v_z)$, with zero mean and standard deviation

$$\sigma_v = \sqrt{\frac{k_BT}{m}} \tag{15.82}$$

The probability distribution of each velocity component is itself a zero-mean Gaussian, with the same standard deviation. From (15.81), we get the distribution for the modulus $v$ of the velocity as

$$f_v(v) = \left(\frac{m}{2\pi k_BT}\right)^{3/2} \exp\left(-\frac{mv^2}{2k_BT}\right) 4\pi v^2. \tag{15.83}$$

which is obtained by explicitly writing out the elementary velocity element $\mathrm{d}^3\vec{v} = 4\pi v^2\mathrm{d}v$. It is easy to show that with this distribution (15.81), the collisional integral in the Boltzmann equation (15.58) vanishes, as it should. Indeed, the one-particle distribution function is then

$$f_1(\vec{p}) = \frac{1}{V}f(\vec{p}) = \frac{1}{V}\left(\frac{1}{2\pi mk_BT}\right)^{3/2} \exp\left(-\frac{p^2}{2mk_BT}\right) \tag{15.84}$$

such that its logarithm is

$$\ln f_1 = -\ln V - \frac{3}{2}\ln\left(2\pi mk_BT\right) - \frac{p^2}{2mk_BT} \tag{15.85}$$

and is indeed a collisional invariant.

### Characteristic velocities

From the Maxwell distribution, we can compute a set of characteristic velocities, the **RMS velocity** $v_q$ (for *Root Mean Square*), the **most probable velocity** $v_p$, and the **mean velocity** $\langle v\rangle$, that are formally defined by

$$v_q^2 = \int v^2 f_v(v)\mathrm{d}v \qquad \frac{\partial f_v}{\partial v}(v_p) = 0 \qquad \langle v\rangle = \int v f_v(v)\mathrm{d}v \tag{15.86}$$

The computation of these quantities for the Maxwell-Boltzmann distribution, using the Gaussian integrals given in an earlier footnote, yields

$$v_q = \sqrt{\frac{3k_BT}{m}} \qquad v_p = \sqrt{\frac{2k_BT}{m}} \qquad \langle v \rangle = \sqrt{\frac{8k_BT}{\pi m}} \tag{15.87}$$

To get an order of magnitude, if we write $\mu = m/m_\mathrm{p}$ with $m_\mathrm{p} = 1.673 \times 10^{-27}\,\mathrm{kg}$ the proton mass, and with $T$ in Kelvin, we have

$$v_q = 0.157\sqrt{\frac{T}{\mu}}\,\mathrm{km\,s^{-1}} \qquad v_p = 0.128\sqrt{\frac{T}{\mu}}\,\mathrm{km\,s^{-1}} \qquad \langle v \rangle = 0.145\sqrt{\frac{T}{\mu}}\,\mathrm{km\,s^{-1}} \tag{15.88}$$

so that, for instance, the typical velocity of hydrogen atoms in the cold neutral interstellar medium (with temperatures of the order 100 K) is about $1\,\mathrm{km\,s^{-1}}$.

## 15.3 Ideal gas equation of state

In the framework of kinetic theory, it is possible to derive an expression for the pressure $P$ based on the distribution of particle velocities. In this model for an ideal monoatomic gas, pressure is computed from the transfer of momentum to the walls of the box in which the gas is held.

### 15.3.1 Kinetic pressure

#### Determination

Consider a particle with initial velocity $\vec{v}$, that is undergoing a collision with a section $\mathrm{d}S$ of the box between $t$ and $t + \mathrm{d}t$. We assume the section to be perpendicular to the $x$ axis, so $\mathrm{d}\vec{S} = \mathrm{d}S\vec{e}_x$. After the collision, the particle has a new velocity $\vec{v'}$, so that its momentum change is $m\Delta\vec{v} = m(\vec{v'} - \vec{v})$ and is given by

$$m\Delta\vec{v} = m(\vec{v'} - \vec{v}) = \int_0^\tau \vec{F}\mathrm{d}t, \tag{15.89}$$

where $\vec{F}$ is the instantaneous force exerted by the box on the particle during the collision, and $\tau$ is the duration of the collision. This force is perpendicular to the wall, so the $y$ and $z$ components of the velocity do not change. Since the momentum $\mathrm{d}^2\vec{p}_\mathrm{box}$ acquired by the wall in that process is the opposite of the particle's momentum change, according to Newton's laws, we have

$$\mathrm{d}^2\vec{p}_\mathrm{box} = 2mv_x\vec{e}_x = 2mv\cos\theta\vec{e}_x \tag{15.90}$$

where $\theta$ is the angle between the normal to the wall and the direction of the incident particle. Between $t$ and $t + \mathrm{d}t$, the number of collisions undergone by the section $\mathrm{d}S$ of the wall on the part of particles having such a velocity $\vec{v}$ is

$$\mathrm{d}N(\vec{v}) = nf(\vec{v})\mathrm{d}^3\vec{v}(\vec{e}_x.\vec{v})\mathrm{d}S\mathrm{d}t, \tag{15.91}$$

where we introduce the density of particles $n$, and the volume $(\vec{e}_x\mathrm{d}S).(\vec{v}\mathrm{d}t) = v\cos\theta\mathrm{d}S\mathrm{d}t$ of the cylinder within which the particle must be in order to hit the section of the boundary in the right time-frame. Now, using the isotropy hypothesis, we have [18]

$$f(\vec{v})\mathrm{d}^3\vec{v} = g(v)v^2\mathrm{d}v\sin\theta\mathrm{d}\theta\mathrm{d}\phi, \tag{15.92}$$

18. Note this is not the same $g$ function as in the determination of the Gaussian form of the Maxwell-Boltzmann distribution.

We therefore have, inserting this form in $\mathrm{d}\vec{p}_{\mathrm{box}} = \mathrm{d}N(\vec{v})\mathrm{d}^2\vec{p}_{\mathrm{box}}$ and integrating over directions [19] $(\theta, \phi)$,

$$\mathrm{d}\vec{p}_{\mathrm{box}} = 2nmv^4 g(v)\mathrm{d}v\mathrm{d}S\mathrm{d}t \times \int_0^{\pi/2} \sin\theta\cos^2\theta\,\mathrm{d}\theta \times \int_0^{2\pi}\mathrm{d}\phi = \frac{4\pi}{3}nmv^4 g(v)\mathrm{d}v\mathrm{d}S\mathrm{d}t. \tag{15.93}$$

Now the probability for a particle to have a velocity modulus $v$ to within $\mathrm{d}v$ is precisely $g(v)\times 4\pi v^2\mathrm{d}v =$

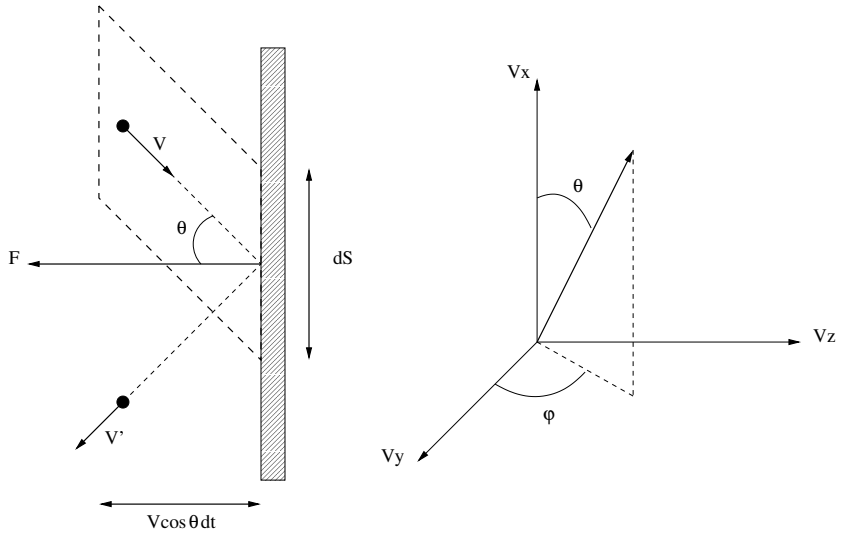



FIGURE 15.4 – Microscopic interpretation of gas pressure.

$f_v(v)\mathrm{d}v$, so the final integration (over $v$) exhibits the mean of $v^2$, i.e. the RMS velocity

$$\mathrm{d}\vec{p}_{\mathrm{box}} = \frac{1}{3}nm\mathrm{d}S\mathrm{d}t \times \int_0^{\infty} v^2 g(v)4\pi v^2\mathrm{d}v = \frac{1}{3}nm\langle v^2\rangle\mathrm{d}S\mathrm{d}t = \frac{1}{3}nmv_q^2\mathrm{d}S\mathrm{d}t \tag{15.94}$$

The pressure is the ratio of the rate of change of the boundary's momentum to the section, so we have

$$P = \frac{|\mathrm{d}\vec{p}_{\mathrm{box}}|}{\mathrm{d}t\mathrm{d}S} = \frac{1}{3}nmv_q^2 \tag{15.95}$$

### A simplistic model

It is possible to find the same expression by considering a simplistic model in which particles have velocities that can only be parallel to one of the $x$, $y$ or $z$ axes, with a single magnitude $u$. In that case, the momentum acquired by the wall in a collision is

$$\mathrm{d}^2\vec{p}_{\mathrm{box}} = 2mu\vec{e}_x \tag{15.96}$$

and the number of such collisions in time $\mathrm{d}t$ is

$$\mathrm{d}N = \frac{1}{6}nu\mathrm{d}S\mathrm{d}t, \tag{15.97}$$

19. Note that integration over $\theta$ is only from $0$ to $\pi/2$ in order to have a collision...

where the factor $1/6$ is to consider only particles with velocity $u\vec{e}_x$. The resulting expression of kinetic pressure is then

$$P = \frac{|\mathrm{d}\vec{p}_{\mathrm{box}}|}{\mathrm{d}t\mathrm{d}S} = \frac{1}{3}nmu^2 \tag{15.98}$$

which is identical to the general one. The fact that the numerical factor $1/3$ is the same should be viewed as a coincidence, however.

### 15.3.2 Kinetic temperature and equation of state

The total microscopic kinetic energy of the gas is the sum of the kinetic energies of the various particles, and it may be written using the RMS velocity $v_q$

$$E_c = \sum_i \frac{1}{2}mv_i^2 = \frac{Nm}{2} \times \frac{1}{N}\sum_i v_i^2 = \frac{Nm}{2}v_q^2 \tag{15.99}$$

The kinetic temperature from the equipartition condition is such that

$$E_c = \frac{3}{2}Nk_BT \tag{15.100}$$

and so we combine these with the equation for the kinetic pressure to obtain the equation of state

$$P = \frac{1}{3}nmv_q^2 = \frac{2E_c}{3V} = \frac{Nk_BT}{V} \tag{15.101}$$

### 15.3.3 A more general expression and an application to white dwarfs

A similar approach may be undertaken to compute the degeneracy pressure of electrons in white dwarfs. The momentum imparted during $\mathrm{d}t$ to a surface $\mathrm{d}S$ by the electrons whose energy is between $\varepsilon$ and $\varepsilon + \mathrm{d}\varepsilon$ is

$$\mathrm{d}\vec{p}_{\mathrm{box}} = n_e(\varepsilon) \times v(\varepsilon)\cos\theta\mathrm{d}t\mathrm{d}S \times 2p(\varepsilon)\cos\theta \times \frac{\mathrm{d}\Omega}{4\pi} \tag{15.102}$$

where[20] $n_e(\varepsilon) = (1/V)\rho(\varepsilon)n_{\mathrm{FD}}(\varepsilon)\mathrm{d}\varepsilon$ is the number of electrons with energies in $[\varepsilon, \varepsilon + \mathrm{d}\varepsilon]$, per unit volume, $v(\varepsilon)\cos\theta\mathrm{d}t\mathrm{d}S$ is the volume of the cylinder where the electrons of appropriate energy where in the preceding time interval in order to hit the wall during $\mathrm{d}t$, $2p(\varepsilon)\cos\theta$ is the momentum imparted to the wall in each of these collisions, and $\mathrm{d}\Omega/(4\pi)$ is the fraction of total solid angle corresponding to the electrons with the appropriate direction. Integrating over that, we have the electronic pressure

$$P_e = \frac{1}{\mathrm{d}S\mathrm{d}t}\int_{\varepsilon_0}^{\infty}\mathrm{d}\varepsilon\int_0^{\pi/2}\mathrm{d}\theta\sin\theta\int_0^{2\pi}\mathrm{d}\phi\left[\frac{\rho(\varepsilon)n_{\mathrm{FD}}(\varepsilon)}{V}v(\varepsilon)\cos\theta\mathrm{d}t\mathrm{d}S \times 2p(\varepsilon)\cos\theta \times \frac{1}{4\pi}\right] \tag{15.103}$$

which may be integrated simply over $\phi$ and $\theta$ to yield the general expression

$$P_e = \frac{1}{3V}\int_{\varepsilon_0}^{\infty}\mathrm{d}\varepsilon\rho(\varepsilon)n_{\mathrm{FD}}(\varepsilon)v(\varepsilon)p(\varepsilon) \tag{15.104}$$

Note that when the temperature is much less than the Fermi temperature ($T \ll T_F$), this becomes

$$P_e = \frac{1}{3V}\int_{\varepsilon_0}^{\varepsilon_{\mathrm{F}}} p(\varepsilon)v(\varepsilon)\rho(\varepsilon)\mathrm{d}\varepsilon. \tag{15.105}$$

20. We recall that $\rho$ is the density of states and $n_{\mathrm{FD}}$ is the Fermi-Dirac distribution for the occupation number.

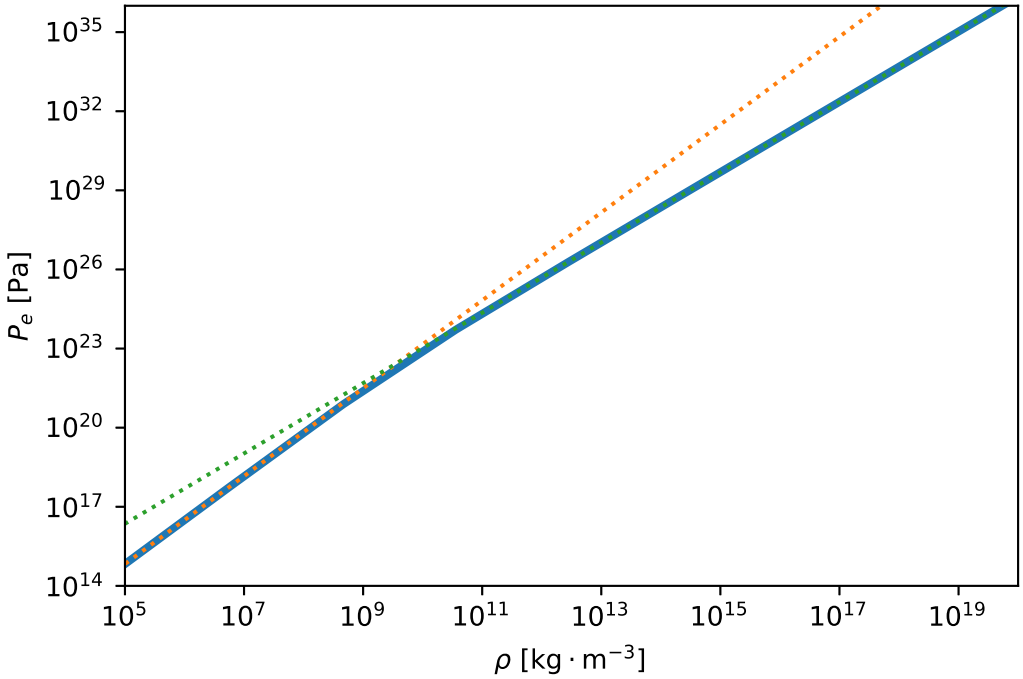



FIGURE 15.5 – Electron degeneracy pressure in a white dwarf as a function of the mass density. The solid line corresponds to the full expression and the dashed lines to the non-relativistic and ultra-relativistic limits.

To go further, we remind that for electrons forming an ideal, relativistic Fermi gas in a box of volume $V$, the density of states is

$$\rho(\varepsilon) = \frac{(2s+1)V}{2\pi^2\hbar^3c^3}(\varepsilon^2 - m_e^2c^4)^{1/2}\varepsilon = \frac{V}{\pi^2\hbar^3c^3}(\varepsilon^2 - m_e^2c^4)^{1/2}\varepsilon \tag{15.106}$$

and that the energy, velocity, and momentum are related by

$$\varepsilon = \gamma m_e c^2 \qquad p = \gamma m_e v \qquad \gamma = \frac{1}{\sqrt{1-\dfrac{v^2}{c^2}}} = \sqrt{1+\frac{p^2}{m_e^2c^2}} \tag{15.107}$$

We then rewrite the expression of the pressure using momentum $p$ as the variable, noting that

$$\sqrt{\varepsilon^2 - m_e^2c^4} = pc \qquad 2c^2p\mathrm{d}p = 2\varepsilon\mathrm{d}\varepsilon \qquad p_\mathrm{F}c = \sqrt{\varepsilon_F^2 - m_e^2c^4} \tag{15.108}$$

This yields

$$P_e = \frac{1}{3V}\int_{m_ec^2}^{\varepsilon_\mathrm{F}} p(\varepsilon)\times\frac{p(\varepsilon)}{\gamma m_e}\times\frac{1}{\pi^2\hbar^3c^3}V\sqrt{\varepsilon^2-m_e^2c^4}\varepsilon\mathrm{d}\varepsilon = \frac{1}{3\pi^2\hbar^3}\int_0^{p_\mathrm{F}}\frac{p^4}{m_e\sqrt{1+\dfrac{p^2}{m_e^2c^2}}}\mathrm{d}p \tag{15.109}$$

Making the final variable change $u = p/(m_e c)$ we get

$$P_e = \frac{m_e^4 c^5}{3\pi^2\hbar^3} F\left(\frac{p_\mathrm{F}}{m_e c}\right) \qquad F(x) = \int_0^x \frac{u^4}{\sqrt{1+u^2}}\mathrm{d}u \tag{15.110}$$

where the function $F$ may be explicited

$$F(x) = \int_0^x \frac{u^4}{\sqrt{1+u^2}}\mathrm{d}u = \frac{1}{8}\left[x(2x^2-3)\sqrt{1+x^2} + 3\ln\left(x+\sqrt{1+x^2}\right)\right].$$

and has asymptotic limits

$$F(x) \simeq \frac{x^5}{5} \quad (x \ll 1) \qquad F(x) \simeq \frac{x^4}{4} \quad (x \gg 1) \tag{15.111}$$

The former limit ($x \ll 1$) corresponds to the non-relativistic regime, the latter ($x \gg 1$) to the ultra-relativistic one. Treating both limits simultaneously by writing $F(x) \simeq x^q/q$,we have

$$P_e \simeq \frac{m_e^4 c^5}{3q\pi^2\hbar^3}\left(\frac{p_\mathrm{F}}{m_e c}\right)^q = \frac{m_e^{4-q}c^{5-q}}{3q\pi^2\hbar^3}\left[\hbar(3\pi^2 n_e)^{1/3}\right]^q = \frac{m_e^{4-q}c^{5-q}}{q(3\pi^2)^{1-q/3}\hbar^{3-q}} n_e^{q/3} \tag{15.112}$$

The density of electrons $n_e$ is related to the mass density of the white dwarf star by

$$n_e = \frac{N_e}{V} = \frac{ZN_n}{V} = \frac{ZM}{Am_p V} = Y_e\frac{\rho}{m_p} \tag{15.113}$$

where $Y_e = Z/A$ is the electronic fraction, ratio of the atomic number to the mass number, $N_n$ is the number of nuclei, and $m_p$ is the mass of a proton. Setting $\alpha = q/3$, we obtain

$$P_e \simeq K_\alpha \rho^\alpha \qquad K_\alpha = \frac{1}{3\alpha}(3\pi^2\hbar^3)^{\alpha-1} m_e^{4-3\alpha} c^{5-3\alpha}\left(\frac{Y_e}{m_p}\right)^\alpha \tag{15.114}$$

We therefore have, in both regimes, a **polytropic equation of state**,

$$P_e \propto \rho^{5/3} \quad \text{(non-relativistic)} \qquad P_e \propto \rho^{4/3} \quad \text{(ultra-relativistic)} \tag{15.115}$$

This can be used in describing the stability of white dwarfs and estimating the maximum mass of such stars that is the **Chandrasekhar mass**. The degeneracy pressure is shown in Fig 15.5 as a function of the mass density of the white dwarf.

## 15.4 Transport coefficients in kinetic theory

The kinetic theory deals not only with gases in thermodynamic equilibrium, but also very importantly with gases *not* in thermodynamic equilibrium. This means using the theory to derive **transport coefficients** (thermal conductivity, viscosity, ...), as we shall now do.

### 15.4.1 Mean free path

In the gas, the **mean free path** $l$ of particles is the average distance covered by a given particle between two collisions. It may be computed simply in the "hard sphere" model, where particles are considered to be rigid spheres of radius $r$. In that case, we study the motion of a particle by considering first that the others are fixed in space. The trajectory of the particle is a series of straight lines in-between

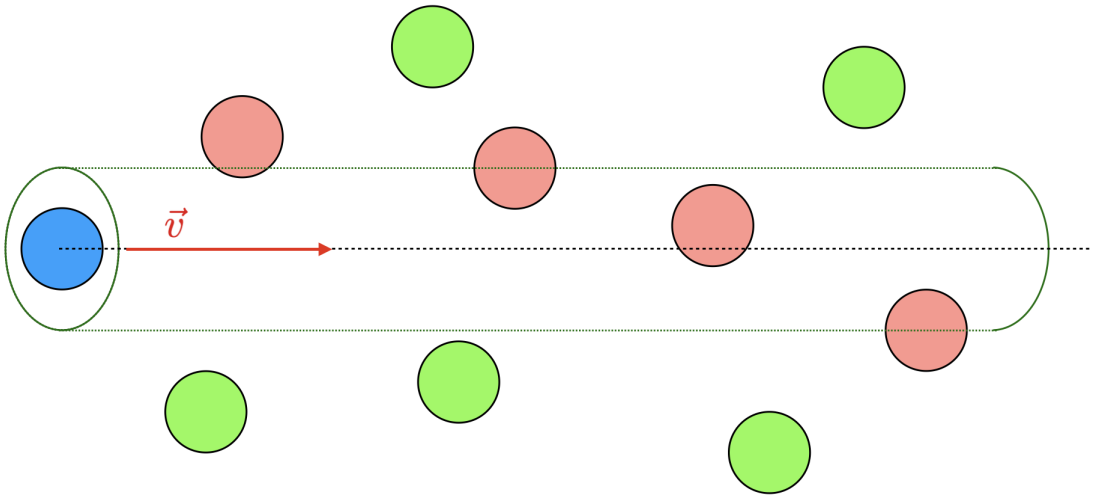



FIGURE 15.6 – Computation of the mean free path. The projectile particle is in blue, the targets that will be hit in red and the ones missed are in green.

collisions, and we represent them aligned, as in Fig. 15.6. The particle collides with any other particle whose center lies within the cylinder of radius $2r$. This means that the **cross-section** is $\sigma = 4\pi r^2$. During time $\mathrm{d}t$, the particle covers an average distance $\langle d\rangle = \langle v\rangle \mathrm{d}t$, so the number of targets hit is $\mathrm{d}N = n\sigma\langle d\rangle = n\sigma\langle v\rangle \mathrm{d}t$. The **mean free time** between two collisions is then

$$\tau = \frac{\mathrm{d}t}{\mathrm{d}N} = \frac{1}{n\sigma\langle v\rangle} \tag{15.116}$$

and the mean free path is $l = \langle v\rangle\tau$, so

$$l = \frac{1}{n\sigma} \tag{15.117}$$

Taking into account the fact that target particles are in fact not fixed, we can still use this result provided we work with the **relative RMS velocity**, which is a factor $\sqrt{2}$ larger[21]. Consequently, the mean free path is lower by the same factor, i.e.,

$$l = \frac{1}{\sqrt{2}n\sigma} \tag{15.118}$$

### 15.4.2 Viscosity

Consider, as in Fig. 15.7, an isothermal planar Couette flow, in which a fluid is located between two plates with different velocities along the $x$ axis. Say that the bottom plate is fixed and the upper plate, at position $y = L$, has velocity $U\vec{e}_x$. The fluid will be entrained so that it has a velocity $\vec{u} = u(y)\vec{e}_x$. At an imaginary horizontal surface $A$ at height $y_A$, the fluid's velocity is $u_0$, and the mean momentum

21. This comes from $\langle(\vec{v}_1 - \vec{v}_2)^2\rangle = \langle\vec{v}_1^2\rangle + \langle\vec{v}_1^2\rangle - 2\langle\vec{v}_1\cdot\vec{v}_2\rangle = 2\langle\vec{v}^2\rangle$, using, in the last step, the decorrelation of the velocities of the two particles.

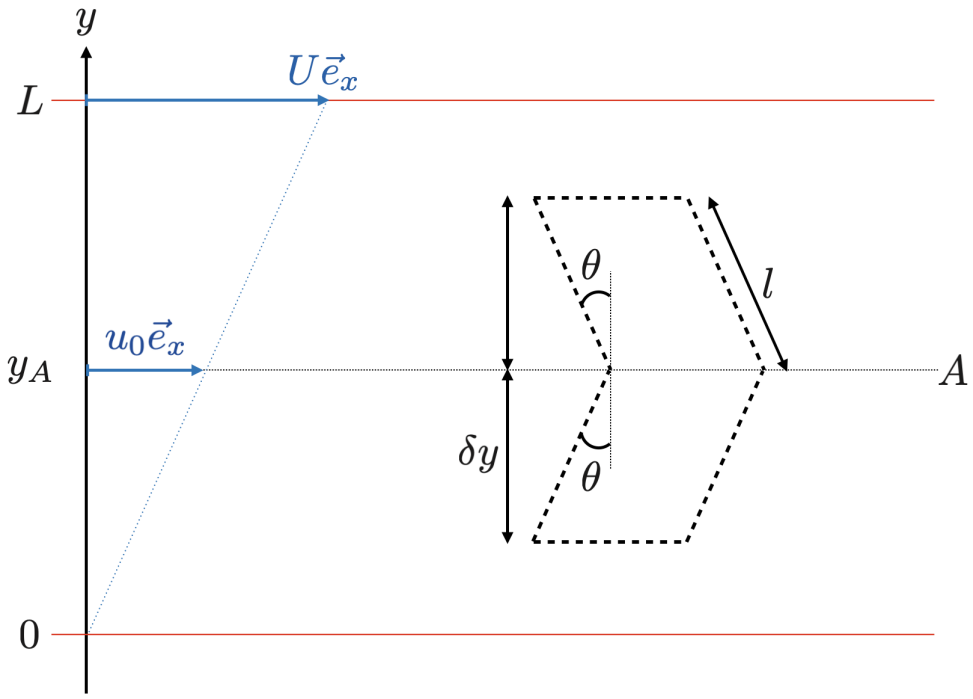



FIGURE 15.7 – Computation of the viscosity coefficient.

$p_x^+$ of the particles above this surface is larger than the mean momentum $p_x^-$ of the particles below it, and the microscopic motion of particles will lead to a transfer of this momentum from top to bottom. Consider the particles above the surface $A$ that cross an element $\mathrm{d}A$ during $\mathrm{d}t$, with velocities in $[v, v+\mathrm{d}v]$ and direction $(\theta, \phi)$. They had their last collision at a height $\delta y = l\cos\theta$ above the surface, so their average forward momentum is

$$p_x^+ = m\left(u_0 + l\cos\theta\frac{\mathrm{d}u}{\mathrm{d}y}\right) \tag{15.119}$$

There are $n \times v\cos\theta \mathrm{d}t\mathrm{d}A \times f(\vec{v})\mathrm{d}^3\vec{v}$ such particles. There are as many that cross the surface coming from below with the same angles $(\theta, \phi)$, but these carry a lower average forward momentum

$$p_x^- = m\left(u_0 - l\cos\theta\frac{\mathrm{d}u}{\mathrm{d}y}\right) \tag{15.120}$$

In total, the net average momentum transfer during $\mathrm{d}t$ across $\mathrm{d}A$ is

$$\mathrm{d}p_x = \int nv\cos\theta\mathrm{d}t\mathrm{d}Af(\vec{v})\mathrm{d}^3\vec{v}\left[m\left(u_0 + l\cos\theta\frac{\mathrm{d}u}{\mathrm{d}y}\right) - m\left(u_0 - l\cos\theta\frac{\mathrm{d}u}{\mathrm{d}y}\right)\right] \tag{15.121}$$

Expliciting the various factors, and noting that $f(\vec{v}) = f_v(v)/(4\pi v^2)$, we have [22]

$$\mathrm{d}p_x = 2nml\mathrm{d}t\mathrm{d}A\frac{\mathrm{d}u}{\mathrm{d}y}\int v\cos^2\theta\frac{f_v(v)}{4\pi}\sin\theta\mathrm{d}\theta\mathrm{d}\phi\mathrm{d}v \tag{15.122}$$

22. Note that $n$ and $\mathrm{d}u/\mathrm{d}y$ are constants.

where the integral covers $v > 0$, $\theta \in [0, \pi/2]$, and $\phi \in [0, 2\pi]$. This yields

$$\mathrm{d}p_x = \frac{1}{3} nml\langle v \rangle \frac{\mathrm{d}u}{\mathrm{d}y} \mathrm{d}t\mathrm{d}A \tag{15.123}$$

and so the net rate of momentum per unit area that is transported across the imaginary surface is

$$\mu = \frac{1}{3} nml\langle v \rangle \frac{\mathrm{d}u}{\mathrm{d}y} \tag{15.124}$$

This is to be compared to Newton's law of viscosity, stating that this net rate is proportional to the shear, with the constant of proportionality being the **(dynamic) viscosity** $\eta$,

$$\mu = \eta \frac{\mathrm{d}u}{\mathrm{d}y} \qquad \eta = \frac{1}{3} nml\langle v \rangle \tag{15.125}$$

For an ideal gas with a Maxwell-Boltzmann distribution of velocities and the above expression for the mean free path, we have

$$\eta = \frac{1}{3} \times m \times \frac{1}{\sqrt{2}\sigma} \times \sqrt{\frac{8k_B T}{\pi m}} = \frac{2}{3\sigma} \sqrt{\frac{mk_B T}{\pi}} \tag{15.126}$$

### 15.4.3 Thermal conductivity

Following the same logic, we can obtain an expression of thermal conductivity, by considering a situation in which some gas is contained between two parallel plates that are maintained at different temperatures $T(y = 0) = T_0$ and $T(y = L) = T_1 > T_0$. Above the imaginary surface $A$, particles have a mean kinetic energy that is larger than that of the particles below the surface. In crossing that surface, they carry this energy with them, leading to a transport of heat. The same arguments developed above lead us to consider the mean kinetic energies $\varepsilon^{\pm}$ of particles that had their last collision a distance $l \cos\theta$ above or below the surface,

$$\varepsilon^{\pm} = \varepsilon_0 \pm \delta\varepsilon = \varepsilon_0 \pm C\delta T = \varepsilon_0 \pm C\delta y \frac{\mathrm{d}T}{\mathrm{d}y} = \varepsilon_0 \pm mc_v l \cos\theta \frac{\mathrm{d}T}{\mathrm{d}y} \tag{15.127}$$

where $C$ is the heat capacity per particle, and $c_v$ is the specific (i.e., per unit mass) heat capacity. The rate of energy transfer per unit surface area, i.e., the heat flux, is then, in absolute value

$$J_Q = \frac{1}{\mathrm{d}A\mathrm{d}t} \times \int nv \cos\theta \mathrm{d}t\mathrm{d}Af(\vec{v})\mathrm{d}^3\vec{v} \times 2mc_v l \cos\theta \frac{\mathrm{d}T}{\mathrm{d}y} \tag{15.128}$$

which gives, expliciting the integrals,

$$J_Q = nmc_v l \frac{\mathrm{d}T}{\mathrm{d}y} \int v \cos^2\theta f_v(v) \sin\theta \mathrm{d}\theta = \frac{1}{3} nmc_v l \frac{\mathrm{d}T}{\mathrm{d}y} \langle v \rangle \tag{15.129}$$

Note that this is a downward flux, considering the setup, so in vectorial form we obtain **Fourier's law** with an expression for the thermal conductivity $K$

$$\vec{J}_Q = -K\vec{\nabla}T \qquad K = \frac{1}{3} nmc_v l\langle v \rangle \tag{15.130}$$

### 15.4.4 Diffusion coefficient

Finally, consider a steady diffusion between two regions of the same gas with perfectly flat and parallel boundaries separated by a layer of the same gas. Both regions have uniform number densities, but the upper region has a higher number density than the lower region. In steady state, the number density at any point is independent of time, but the density $n$ increases uniformly with distance $y$ from the lower plate. The particles that cross the imaginary surface $A$ during $\mathrm{d}t$ from above have a larger density than those crossing the same surface in the same time coming from below. The corresponding densities are

$$n^{\pm} = n_0 \pm l\cos\theta\frac{\mathrm{d}n}{\mathrm{d}y} \tag{15.131}$$

where $n_0$ is the density of particles at the surface $A$. Once again [23] counting the particles per element of velocity and integrating, we obtain the net number of particles crossing per unit time and unit surface, i.e., the particle density current, as

$$J_n = \frac{\mathrm{d}^2N}{\mathrm{d}A\mathrm{d}t} = \int v\cos\theta f(\vec{v})\mathrm{d}^3\vec{v}\times 2l\cos\theta\frac{\mathrm{d}n}{\mathrm{d}y} \tag{15.132}$$

which yields, once again expliciting the integrals,

$$J_n = \frac{1}{3}l\langle v\rangle\frac{\mathrm{d}n}{\mathrm{d}y} \tag{15.133}$$

Note that this is yet again a downward flux, considering the setup, so in vectorial form we obtain **Fick's law** with an expression for the diffusion coefficient $D$

$$\vec{J_n} = -D\vec{\nabla}n \qquad D = \frac{1}{3}l\langle v\rangle \tag{15.134}$$

## 15.5 From kinetic theory to hydrodynamics

At equilibrium, the Boltzmann equation's solution is, as we saw, the Maxwell-Boltzmann distribution. One may wonder if non-equilibrium dynamics, i.e., the equations of fluid dynamics such as Euler or Navier-Stokes equations, could be obtained from the kinetic theory as well.

### 15.5.1 Equations of fluid dynamics

#### Principle

The equations of fluid dynamics may be obtained from the Boltzmann equation, by forming moments with quantities of the type $Q(\vec{q},\vec{p})$, i.e.,

$$\int Q(\vec{q},\vec{p})\left[\frac{\partial f_1}{\partial t}+\frac{\vec{p}_1}{m}\cdot\frac{\partial f_1}{\partial\vec{q}_1}-\frac{\partial U_{\mathrm{ext}}}{\partial\vec{q}_1}\cdot\frac{\partial f_1}{\partial\vec{p}_1}\right]\mathrm{d}^3\vec{p} = \int Q(\vec{q},\vec{p})\left(\frac{Df_1}{Dt}\right)_{\mathrm{coll}}\mathrm{d}^3\vec{p} \tag{15.135}$$

If $Q$ is a collisional invariant, the right-hand side of this equation is null [24]. If furthermore $Q$ does not explicitly depend on position $\vec{q}$, and the forces on the particles do not depend on their velocities, then this leads to a conservation equation

$$\frac{\partial}{\partial t}\left(n\langle Q\rangle\right)+\sum_i\frac{1}{m}\frac{\partial}{\partial q_i}\left(n\langle Qp_i\rangle\right)-n\sum_i\dot{p}_i\left\langle\frac{\partial Q}{\partial p_i}\right\rangle = 0 \tag{15.136}$$

23. One should be careful here that the spatial factor in the distribution function is not $n$, but should be taken as $n^+$ or $n^-$ depending on the side of the surface that is considered. This is why the $n$ factor under the integral is not present.
24. See "Foundations of radiation hydrodynamics" (Mihalas & Mihalas) for a demonstration.

where $i$ runs from 1 to 3 for the three spatial coordinates, $n$ is the number density of particles

$$n = \int f_1 \left(\vec{r}, \vec{p}, t\right) \mathrm{d}^3\vec{p} \tag{15.137}$$

and the spatial averages are of course taken over the distribution function, e.g.,

$$\langle Q \rangle = \int Q f_1 \mathrm{d}^3\vec{q} \mathrm{d}^3\vec{p} \tag{15.138}$$

By successively taking $Q = m$, $Q = mv_i$, and $Q = mv^2/2$, which are collisional invariants, this conservation theorem yields the equations of fluid dynamics

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{u}) = 0 \tag{15.139}$$

$$\rho \left[ \frac{\partial \vec{u}}{\partial t} + \left( \vec{u} \cdot \vec{\nabla} \right) \vec{u} \right] = -\vec{\nabla} p + \vec{f} + \vec{\nabla} \cdot \vec{\sigma} \tag{15.140}$$

$$\frac{\partial}{\partial t} \left[ \rho \left( e + \frac{1}{2} u^2 \right) \right] + \vec{\nabla} \cdot \left[ \rho \left( h + \frac{1}{2} u^2 \right) - \vec{u} \cdot \vec{\sigma} + \vec{J_Q} \right] = \vec{f} \cdot \vec{v} \tag{15.141}$$

In these equations, $\rho = mn$ is the mass density, $\vec{u}$ is the local mean speed, $\vec{f}$ is the externally-applied force per unit volume, $e$ is the specific internal energy (i.e., by unit mass), $\vec{\sigma}$ is the viscous stress tensor, and $\vec{J_Q}$ is the heat flux. These last two quantities depend on the single-particle distribution $f_1$, because they are expressed as

$$\vec{J_Q} = \frac{1}{2} \rho \left\langle (\delta u)^2 \, \vec{\delta u} \right\rangle \qquad \sigma_{ij} = -\rho \left[ \langle \delta u_i \delta u_j \rangle - \frac{1}{3} \langle (\, \delta u \,)^2 \delta_{ij} \rangle \right] \tag{15.142}$$

where $\vec{\delta u} = \vec{v} - \vec{u}$ is the particle velocity, once the average fluid velocity is removed. Symmetry considerations show that at global equilibrium, so when $f_1$ takes the Maxwellian form $f_1^0$ with parameters that do not depend on position and time, then $\vec{J}_Q^0 = \vec{0}$ and $\vec{\sigma}^0 = \vec{0}$. They can be non-null only when out of equilibrium.

### Example demonstration

For instance, let us demonstrate the first of these equations, which is simply the conservation of mass. By taking $Q = m$, Eq. (15.136) becomes

$$\frac{\partial}{\partial t} \left( n \langle m \rangle \right) + \sum_i \frac{1}{m} \frac{\partial}{\partial q_i} \left( n \langle m p_i \rangle \right) - n \sum_i \dot{p}_i \left\langle \frac{\partial m}{\partial p_i} \right\rangle = 0 \tag{15.143}$$

Of course, the mass does not depend on any coordinate, so $\langle m \rangle = m$, and the last sum is necessarily zero. Inserting the mass density $\rho = nm$, we have

$$\frac{\partial \rho}{\partial t} + \sum_i \frac{1}{m} \frac{\partial}{\partial q_i} \left( n \langle m p_i \rangle \right) = 0 \tag{15.144}$$

In the remaining sum over the three coordinates $i = 1, 2, 3$, we have

$$\frac{1}{m} \frac{\partial}{\partial q_i} \left( n \langle m p_i \rangle \right) = \frac{\partial}{\partial q_i} \left( n \langle p_i \rangle \right) = \frac{\partial}{\partial q_i} \left( n \langle m v_i \rangle \right) = \frac{\partial}{\partial q_i} \left( \rho \langle v_i \rangle \right) = \frac{\partial}{\partial q_i} \left( \rho u_i \right) \tag{15.145}$$

so the sum of these three terms is precisely the divergence of $\rho \vec{u}$,

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{u}) = 0 \tag{15.146}$$

### 15.5.2 Chapman-Enskog development and transport coefficients

The purpose of the **Chapman-Enskog** theory is to develop a perturbative approach to small deviations from the Maxwell distribution, for systems in which a slight imbalance is present (for instance, a small temperature gradient). In doing so, expressions for transport coefficients are obtained in terms of molecular data. In this perturbative approach, the small parameter controlling the development is the **Knudsen number**

$$\xi = \frac{l}{L} \tag{15.147}$$

which is the ratio of the mean free path of particles, $l$, to the characteristic scale of the flow, $L$. The Chapman-Enskog development seeks successive approximations to $f_1$, starting from the Maxwellian $f_1^0$. First, the Boltzmann equation is written formally as

$$\xi \mathfrak{D} f_1 = J\left(f_1, f_1\right) \tag{15.148}$$

where $\mathfrak{D}$ is the differential operator acting on the left-hand side of the Boltzmann equation, and $J$ is a functional representing the collisional integral. The arguments of this functional involve integrals over terms of the form[25] $g'h' - gh$. In the limit $\xi \to 0$ one should recover the Maxwell-Boltzmann distribution, and so the development is in powers of $\xi$ for $\xi \ll 1$,

$$f_1 = f_1^{(0)} + \xi f_1^{(1)} + \xi^2 f_1^{(2)} + \dots \tag{15.149}$$

where $f_1^{(0)}$ is the Maxwell-Boltzmann distribution. Inserting this development in the Boltzmann equation, we obtain, at successive orders,

$$J\left(f_1^{(0)}, f_1^{(0)}\right) = 0 \qquad \mathfrak{D} f_1^{(0)} = J\left(f_1^{(1)}, f_1^{(0)}\right) + J\left(f_1^{(0)}, f_1^{(1)}\right) \qquad \dots \tag{15.150}$$

The first of these equations yields the Maxwellian form of $f_1^{(0)}$, while the second one implicitly gives the second term of the Chapman-Enskog development, $f_1^{(1)}$. It is well beyond the scope of this course to proceed with these calculations[26], but the result is[27]

$$f_1^{(1)} = \left[-\frac{1}{n}\sqrt{\frac{2k_BT}{m}}\vec{A}(\vec{v}) \cdot \vec{\nabla}\left(\ln T\right) - \frac{2}{n}\mathbb{B}(\vec{v}) : \vec{\nabla} \otimes \vec{u}\right] f_1^{(0)} \tag{15.151}$$

From this, one gets expressions of the heat flux $\vec{J}_Q$ and momentum-flux tensor $\sigma_{ij}$ as

$$\vec{J}_Q = -K\vec{\nabla}T \qquad \sigma_{ij} = P\delta_{ij} - \eta\left[\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) - \frac{2}{3}\left(\vec{\nabla} \cdot \vec{u}\right)\delta_{ij}\right] \tag{15.152}$$

where $K$ and $\eta$ appear as, respectively, the thermal conductivity and the dynamic viscosity coefficient, and the method yields expressions for these. One important prediction is that the viscosity is independent of the density, a result that is well in agreement with experiments. On the other hand, it does depend on temperature, and so does the thermal conductivity. For instance, in the hard-sphere model with radius $r$, we have

$$\eta = 1.016\frac{5}{64r^2}\sqrt{\frac{mk_BT}{\pi}} \qquad K = 2.522\frac{3k_B}{2m}\eta \tag{15.153}$$

25. Primed quantities refer to post-collision states, unprimed ones to pre-collision ones.
26. See Chapman & Cowling (1970), *The Mathematical Theory of Non-Uniform Gases*, Cambridge University Press.
27. $\vec{A}(\vec{v})$ is a vector and $\mathbb{B}(\vec{v})$ is a rank-two tensor, like $\vec{\nabla} \otimes \vec{u}$ with which it is combined using a double scalar product.

and such a temperature dependence is also apparent in other interaction models. When the interaction potential between particles is $\Phi(r) \propto r^{-\alpha}$, both of these coefficients vary with temperature as $T^s$, with an exponent

$$s = \frac{1}{2} + \frac{2}{\alpha} \qquad \alpha = \frac{4}{2s-1} \tag{15.154}$$

In the hard sphere model, $\alpha \to \infty$, leading indeed to $K \propto \sqrt{T}$ and $\eta \propto \sqrt{T}$. Conversely, we can use the observed scaling of the transport coefficients with temperature to constrain the interaction potential. A reasonable agreement with experiments is obtained for $s \approx 0.668$, corresponding to $\alpha \approx 11.9$. This is a justification of the short-range repulsion exponent usually taken for the **Lennard-Jones** potential.

# 16

# Introduction to stochastic processes : random walks

The topic of **stochastic processes** includes a number of various problems in which randomness is of the essence. Chief among these is the paradigm of **random walks**, with applications to polymers, domain growth, voting problems, diffusion-reaction problems, percolation, ..., from which it is possible to recover diffusion equations.

## 16.1 Jump probabilities

### 16.1.1 Definition

We consider a $d$-dimensional hypercubic regular **lattice**, as shown in Fig. 16.1 for the case $d = 2$. We write $\Omega(\vec{x}, t)$ for the number of paths going from the origin $\vec{0}$ to $\vec{x}$ in $t$ steps. Note that both spatial and temporal variables are discrete. At each step, there are $2d$ branching possibilities, so the total number of paths of length $t$ is [1]

$$\sum_{\vec{x}} \Omega(\vec{x}, t) = (2d)^t \tag{16.1}$$

Thus, the probability $P(\vec{x}, t)$ of finding the walker at $\vec{x}$ at time $t$ is

$$P(\vec{x}, t) = \frac{\Omega(\vec{x}, t)}{(2d)^t} \tag{16.2}$$

The time evolution of this probability will constitute the **master equation**. For an **isotropic walk**, there is an equal probability for a walker to arrive at $\vec{x}$ at time $t + 1$ coming from any of the adjacent sites $\vec{x} + \vec{e}_i$ $(i = 1, \ldots, 2d)$ at the previous time step, so [2]

$$P(\vec{x}, t+1) = \sum_{i=1}^{2d} \frac{1}{2d} P(\vec{x} + \vec{e}_i) \tag{16.3}$$

where the factor $1/(2d)$ is the **transition probability** that the walker will go to $\vec{x}$ rather than to any other site adjacent to $\vec{x} + \vec{e}_i$ in the time step $t \to t + 1$.

1. We include self-intersecting walks, the case of self-avoiding walks being much more complex.
2. Note that in the notation $\{\vec{e}_i\}$ with $i = 1, \ldots, 2d$, we distinguish between $\vec{e}_x$ and $-\vec{e}_x$.

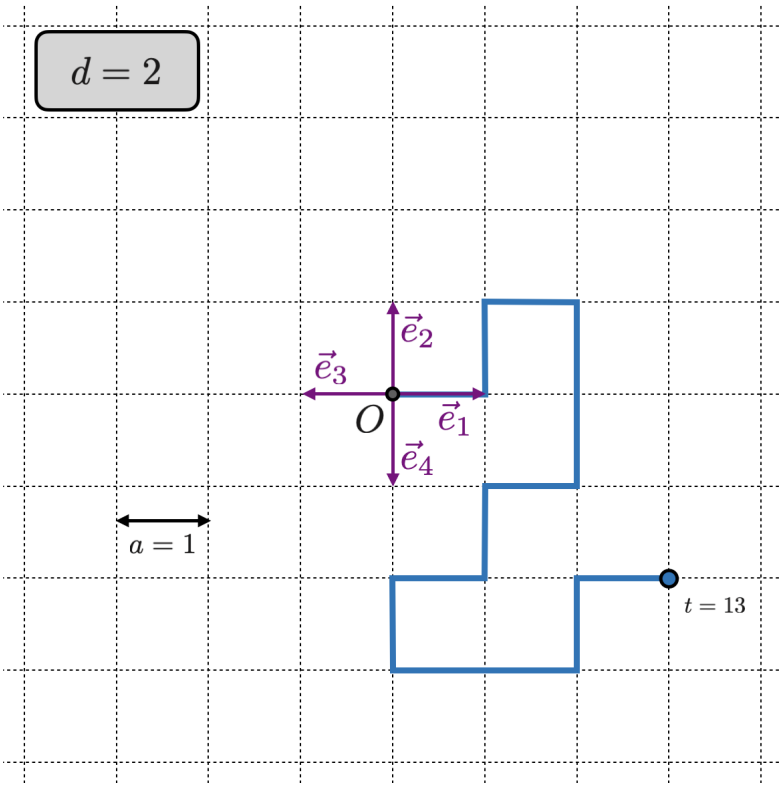



FIGURE 16.1 – A $d = 2$ regular lattice with step $a = 1$, showing the unit vectors for all four directions, and a random walker's path from $t = 0$ to $t = 13$.

### 16.1.2 Computation of the probabilities in the $d = 1$ case

In the case $d = 1$, Eq. (16.3) becomes

$$P(x, t+1) = \frac{1}{2}\left[P(x-1, t) + P(x+1, t)\right] \tag{16.4}$$

In the left and right sides of this equation, both variables of $P$ change, which makes it difficult to solve. To overcome this problem, we introduce the **generative function**

$$Q(\lambda, t) = \sum_{x=-\infty}^{+\infty} \lambda^x P(x, t) \tag{16.5}$$

Obviously, the equation on $P$ yields the following relationship for $Q$,

$$Q(\lambda, t+1) = \frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right) Q(\lambda, t) \tag{16.6}$$

which makes it much easier to compute, provided we know the initial value $Q(\lambda, 0)$. Let us assume that the initial position of the walker is perfectly known, and choose that position as the origin, i.e., $P(x, 0) = \delta_{x,0}$ where $\delta_{i,j}$ is the **Kronecker delta**. Then $Q(\lambda, 0) = 1$ and

$$Q(\lambda, t) = \left[\frac{1}{2}\left(\lambda + \frac{1}{\lambda}\right)\right]^t = \frac{1}{2^t}\sum_{k=0}^{t} C_t^k \lambda^{t-2k} = \frac{1}{2^t}\left(\lambda^t + C_t^1 \lambda^{t-2} + \ldots + C_t^{t-1}\lambda^{2-t} + \lambda^{-t}\right) \tag{16.7}$$

The identity of this polynomial in $\lambda$ with that in the definition (16.5) of $Q$ yields that, in this case of a start at the origin,

— If $|x| > t$, then $P(x,t) = 0$
— If $|x| \leqslant t$ and $t - x$ is odd, then $P(x,t) = 0$
— If $|x| \leqslant t$ and $t - x$ is even, then $P(x,t) = 2^{-t} C_t^{(t-x)/2}$

### 16.1.3 Computation of the probabilities in the $d = 2$ case

In the two-dimensional case $d = 2$, we write $\vec{x} = (x_1, x_2)$ and the master equation (16.3) becomes

$$P(x_1, x_2, t+1) = \frac{1}{4} \left[ P(x_1 - 1, x_2, t) + P(x_1 + 1, x_2, t) + P(x_1, x_2 - 1, t) + P(x_1, x_2 + 1, t) \right] \quad (16.8)$$

Assuming again that we start from the origin $P(x_1, x_2, 0) = \delta_{x_1,0} \delta_{x_2,0}$, and introducting the two-dimensional generative function

$$Q(\lambda_1, \lambda_2, t) = \sum_{x_1} \sum_{x_2} \lambda_1^{x_1} \lambda_2^{x_2} P(x_1, x_2, t) \quad (16.9)$$

where again $x_1$ and $x_2$ extend a priori from $-\infty$ to $+\infty$, we have

$$Q(\lambda_1, \lambda_2, t+1) = \frac{1}{4} \left[ \lambda_1 + \frac{1}{\lambda_1} + \lambda_2 + \frac{1}{\lambda_2} \right] Q(\lambda_1, \lambda_2, t) = \frac{1}{4^t} \left( \lambda_1 + \frac{1}{\lambda_1} + \lambda_2 + \frac{1}{\lambda_2} \right)^t \quad (16.10)$$

where the last equation only holds because we start from the origin.

### 16.1.4 Computation of the probabilities in the general case

For any dimension $d$, we can generalize the above results by introducing the probability $P(x_1, \ldots, x_d, t)$ and the following generative function as a $d$-dimensional sum

$$Q(\lambda_1, \ldots, \lambda_d, t) = \sum_{x_1} \ldots \sum_{x_d} \lambda_1^{x_1} \ldots \lambda_d^{x_d} P(x_1, \ldots, x_d, t) \quad (16.11)$$

Assuming that the walker is at the origin initially with 100% probability, we have

$$Q(\lambda_1, \ldots, \lambda_d, t) = \frac{1}{(2d)^t} \left( \lambda_1 + \frac{1}{\lambda_1} + \ldots + \lambda_d + \frac{1}{\lambda_d} \right)^t \quad (16.12)$$

To recover the probabilities from the generative functions, we note that this is a general inversion problem, aiming to recover the function $a(x)$ from the function $f(\lambda)$ defined as

$$f(\lambda) = \sum_x \lambda^x a(x) \quad (16.13)$$

Note that this is one-dimensional but will be generalized later. Writing this for $\lambda = e^{i\theta}$ we have

$$f\left(e^{i\theta}\right) = \sum_x e^{ix\theta} a(x) \quad (16.14)$$

which is a Fourier series. In this form $a(x)$ appears as a Fourier coefficient of $g(\theta) = f\left(e^{i\theta}\right)$ and can thus be recovered by the following relation from Fourier analysis [3]

$$a(x) = \frac{1}{2\pi} \int_0^{2\pi} e^{-ix\theta} f\left(e^{i\theta}\right) \mathrm{d}\theta \quad (16.15)$$

3. See Pinkus & Zafrany, *Fourier Series and Integral Transforms*, Cambridge University Press (1997), p. 42-44.

We generalize this to $d$ dimensions using the expression of the generative function above as

$$P(\vec{x},t)=\int_0^{2\pi}\frac{\mathrm{d}\theta_1}{2\pi}\ldots\int_0^{2\pi}\frac{\mathrm{d}\theta_d}{2\pi}e^{-i\vec{x}\cdot\vec{\theta}}\left(\frac{e^{i\theta_1}+e^{-i\theta_1}+\cdots+e^{i\theta_d}+e^{-i\theta_d}}{2d}\right)^t \tag{16.16}$$

where we have introduced the vector $\vec{\theta}=(\theta_1,\ldots,\theta_d)$. This can also be written in the equivalent form

$$P(\vec{x},t)=\int_0^{2\pi}\frac{\mathrm{d}\theta_1}{2\pi}\ldots\int_0^{2\pi}\frac{\mathrm{d}\theta_d}{2\pi}e^{-i\vec{x}\cdot\vec{\theta}}\left(\frac{1}{d}\sum_{i=1}^{d}\cos\theta_i\right)^t \tag{16.17}$$

### 16.1.5 Asymptotic behaviour at long times

One of the main questions in studies of random walks is to find the probability to reach the origin, in which time, and if so how many times. This involves the computation of the asymptotic behaviour of $P(\vec{0},t)$ in the limit $t\gg 1$.

For $d=1$, this probability is zero for odd time steps $t$, and if $t$ is even it is

$$P(0,t)=\frac{1}{2^t}C_t^{t/2} \tag{16.18}$$

an expression that can be developed, using the Stirling approximation, into [4]

$$P(0,t)\approx\frac{1}{\sqrt{2\pi t}}\left[1+(-1)^t\right] \tag{16.20}$$

In the general $d$-dimensional case, we have

$$P(\vec{0},t)\approx\frac{d^{d/2}}{(2\pi t)^{d/2}}\left[1+(-1)^t\right] \tag{16.21}$$

This can be shown by writing the probability in full from Eq. (16.17)

$$P(\vec{0},t)=\frac{1}{(2\pi)^d}\int_0^{2\pi}\mathrm{d}\theta_1\ldots\int_0^{2\pi}\mathrm{d}\theta_d\left(\frac{\cos\theta_1+\ldots\cos\theta_d}{d}\right)^t \tag{16.22}$$

and noting that in the limit $t\to\infty$, the integrand vanishes, except in the vicinity of $\theta_1=\ldots=\theta_d=0$ or $\theta_1=\ldots=\theta_d=\pi$, because in any other configuration, the sum of cosines is $<d$. Writing the development of the cosines in the vicinity of zero yields the first term of Eq. (16.21), and writing it in the vicinity of $\pi$ yields the second [5].

4. Indeed, we have

$$\frac{1}{2^t}C_t^{t/2}=\frac{t!}{2^t[(t/2)!]^2}\approx\frac{1}{2^t}\frac{\left(\frac{t}{e}\right)^t\sqrt{2\pi t}}{\left[\left(\frac{t}{2e}\right)^{t/2}\sqrt{\pi t}\right]^2}=\sqrt{\frac{2}{\pi t}} \tag{16.19}$$

and the factor $\left[1+(-1)^t\right]/2$ takes care of the distinction between odd and even $t$ values.

5. In the vicinity of $\theta_1=\ldots=\theta_d=0$, we have

$$P(\vec{0},t)\approx\frac{1}{(2\pi)^d}\int_0^{2\pi}\mathrm{d}\theta_1\ldots\int_0^{2\pi}\mathrm{d}\theta_d\left(\frac{1-\frac{\theta_1^2}{2}+\ldots+1-\frac{\theta_d^2}{2}}{d}\right)^t \tag{16.23}$$

which is also the development of an exponential integral

$$P(\vec{0},t)\approx\frac{1}{(2\pi)^d}\int_0^{2\pi}\mathrm{d}\theta_1\ldots\int_0^{2\pi}\mathrm{d}\theta_d\exp\left[-t\left(\frac{\theta_1^2}{2d}+\ldots+\frac{\theta_d^2}{2d}\right)\right]\approx\frac{d^{d/2}}{(2\pi t)^{d/2}} \tag{16.24}$$

In the vicinity of $\theta_1=\ldots=\theta_d=\pi$, the same result is obtained with a factor $(-1)^t$.

## 16.2 Moments of the position

### 16.2.1 One-dimensional case

By definition, the **moments** of the position of the walker are, in one dimension ($d = 1$)

$$\langle x^n \rangle(t) = \sum_x x^n P(x,t) \tag{16.25}$$

By using the definition of the generating function, we have

$$\lambda \frac{\partial Q}{\partial \lambda} = \sum_x x \lambda^x P(x,t) \tag{16.26}$$

and by a successive application of the $\lambda(\partial/\partial\lambda)$ operator, we obtain

$$\left(\lambda \frac{\partial}{\partial \lambda}\right)^n Q(\lambda,t) = \sum_x x^n \lambda^x P(x,t) \tag{16.27}$$

so that the $n^{\text{th}}$ moment of the position is obtained by taking the value of this function for $\lambda = 1$, i.e.,

$$\langle x^n \rangle(t) = \left(\lambda \frac{\partial}{\partial \lambda}\right)^n Q(\lambda,t)_{|\lambda=1} \tag{16.28}$$

This gives that the first moment, i.e. the average position, is zero, $\langle x \rangle(t) = 0$, as can be demonstrated

$$\lambda \frac{\partial Q}{\partial \lambda} = \lambda \frac{t}{2^t} \left(\lambda + \frac{1}{\lambda}\right)^{t-1} \left(1 - \frac{1}{\lambda^2}\right) = \frac{t}{2^t} \left(\lambda - \frac{1}{\lambda}\right) \left(\lambda + \frac{1}{\lambda}\right)^{t-1} \tag{16.29}$$

so that, taking this expression for $\lambda = 1$ indeed yields

$$\langle x \rangle(t) = 0 \tag{16.30}$$

which is of course not surprising for an unbiased random walk that starts from the origin. The following moments are

$$\langle x^2 \rangle(t) = t \qquad \langle x^3 \rangle(t) = 0 \qquad \langle x^4 \rangle(t) = 3t^2 - 2t \tag{16.31}$$

Note that the general expression for the moments is [6]

$$\langle x^{2n+1} \rangle(t) = 0 \qquad \langle x^{2n} \rangle(t) = \frac{(2n)!}{2^n n!} t^n + \mathcal{O}(t^{n-1}) \tag{16.32}$$

Note that we have another way to compute the second-order moment (the variance) as a function of time by making use of the master equation. Indeed,

$$\langle x^2 \rangle(t+1) = \sum_x x^2 P(x,t+1) = \frac{1}{2} \sum_x x^2 \left[P(x-1,t) + P(x+1,t)\right] \tag{16.33}$$

and noting that $x^2 = (x+1)^2 - 2(x+1) + 1 = (x-1)^2 + 2(x-1) + 1$, we have, by judiciously replacing $x^2$ in each of the sums above,

$$\langle x^2 \rangle(t+1) = \frac{1}{2}\left(\langle x^2 \rangle(t) - 2\langle x \rangle(t) + 1 + \langle x^2 \rangle(t) + 2\langle x \rangle(t) + 1\right) = \langle x^2 \rangle(t) + 1 \tag{16.34}$$

Since $\langle x^2 \rangle(t=0) = 0$ when starting from the origin, we recover that $\langle x^2 \rangle(t) = t$.

6. This can be shown by writing $\lambda = e^u$ so that $Q(\lambda,t) = e^{t \ln \cosh u} \approx e^{tu^2/2}$, where the final approximation is valid in the vicinity of $\lambda = 1$, i.e., $u = 0$.

### 16.2.2 General case

For $d > 1$ the general definition of the moments is the ensemble average $\langle x_1^{n_1} \dots x_d^{n_d} \rangle$, which may be shown to be expressed in terms of a specific derivative of the generating function

$$\langle x_1^{n_1} \dots x_d^{n_d} \rangle = \left( \lambda_1 \frac{\partial}{\partial \lambda_1} \right)^{n_1} \dots \left( \lambda_d \frac{\partial}{\partial \lambda_d} \right)^{n_d} Q(\lambda_1, \dots, \lambda_d, t)_{|\lambda_1 = \dots = \lambda_d = 1} \tag{16.35}$$

## 16.3 Visiting the origin

### 16.3.1 Average number of visits to the origin

Between times 0 and $t$, there are $(2d)^t$ possible walks $W$ of $t$ steps. We write $n_t(W)$ to stand for the number of times a specific walk $W$ passes through the origin. Then, the average number of visits to the origin is

$$\langle n_t \rangle = \frac{1}{(2d)^t} \sum_W n_t(W) \tag{16.36}$$

To compute $n_t(W)$, we introduce $g(W, t_1)$ that takes the value $g(W, t_1) = 1$ if the walk $W$ does reach the origin at precisely time $t_1$ and $g(W, t_1) = 0$ otherwise. Obviously,

$$n_t(W) = \sum_{t_1=0}^{t} g(W, t_1) \tag{16.37}$$

and therefore the average number of visits to the origin is

$$\langle n_t \rangle = \frac{1}{(2d)^t} \sum_W \sum_{t_1=0}^{t} g(W, t_1) = \sum_{t_1=0}^{t} \left[ \sum_W \frac{1}{(2d)^t} g(W, t_1) \right] \tag{16.38}$$

By interverting the sum signs, we recognize, in the innermost sum, the probability to reach the origin at time $t_1$, so that

$$\langle n_t \rangle = \sum_{t_1=0}^{t} P(\vec{0}, t_1) \tag{16.39}$$

How does that average number change as a function of dimension ? We consider the asymptotic behaviour in the limit $t \to \infty$, showing that the generic term of the sum (16.39) above is of the order

$$P(\vec{0}, t_1) \approx \frac{d^{d/2}}{(2\pi t_1)^{d/2}} \left[ 1 + (-1)^{t_1} \right] \tag{16.40}$$

when $t_1 \to \infty$. In the limit $t \to \infty$, the generic term of the series representing $\langle n_t \rangle$ is of order $t_1^{-d/2}$, which leads to the following conclusion :

— If $d > 2$, $\langle n_t \rangle$ has a finite limit for $t \to \infty$, so the walk visits the origin a finite number of times.

— If $d \leqslant 2$, $\langle n_t \rangle \to +\infty$, and the walk returns to the origin an infinite number of times.

We interpret this by saying that if the walk is restricted to a space of dimension $d = 1$ or $d = 2$, there is not enough "space" for the walk to avoid the origin, and it is bound to return to the origin an infinite number of times. We say that $d = 2$ is a **critical dimension**, separating two very different behaviours.

### 16.3.2 Reaching the origin from any given point

#### Determination of the probability to reach the origin

So far, our walks started from the origin, but we may ask what is the probability to reach the origin when starting from any point $\vec{x}$. We write $\psi(\vec{x}, t)$ for the probability to have reached the origin at least

once between $0$ and $t$, starting from position $\vec{x}$. Since reaching the origin in $t+1$ time steps is equivalent to reaching the origin from any neighbouring point of $\vec{x}$ in $t$ time steps, we have

$$\psi(\vec{x}, t+1) = \frac{1}{2d} \sum_{i=1}^{2d} \psi(\vec{x} + \vec{e}_i, t) \tag{16.41}$$

with obviously[7] $\psi(\vec{0}, t) = 1$. Our goal is to find the probability to pass through the origin at least once starting from $\vec{x}$, which we write

$$q(\vec{x}) = \lim_{t \to \infty} \psi(\vec{x}, t) \tag{16.42}$$

Taking the limit $t \to \infty$ of the above equation (16.41) we have

$$q(\vec{x}) = \frac{1}{2d} \sum_{i=1}^{2d} q(\vec{x} + \vec{e}_i) \tag{16.43}$$

with obviously $q(\vec{0}) = 1$. From the isotropy property, the value of $q$ for each node that neighbours the origin is independent of the node, so that $q(\vec{e}_i) = a$. Using this, we may rewrite the expression of $q$ as

$$q(\vec{x}) = (1-a)\delta_{\vec{x},\vec{0}} + \frac{1}{2d} \sum_{i=1}^{2d} q(\vec{x} + \vec{e}_i) \tag{16.44}$$

Since this is an equation similar in form to Eq. (16.3), we search for $q$ as an integral of the form of Eq. (16.17), i.e., we look for a function $Q(\theta_1, \ldots, \theta_d)$ such that

$$q(\vec{x}) = \frac{1}{(2\pi)^d} \int_0^{2\pi} \mathrm{d}\theta_1 \ldots \int_0^{2\pi} \mathrm{d}\theta_d Q(\theta_1, \ldots, \theta_d)\, e^{-i\vec{x}\cdot\vec{\theta}} \tag{16.45}$$

but the properties of the Fourier transform state that

$$\delta_{\vec{x},\vec{0}} = \frac{1}{(2\pi)^d} \int_0^{2\pi} \mathrm{d}\theta_1 \ldots \int_0^{2\pi} \mathrm{d}\theta_d e^{-i\vec{x}\cdot\vec{\theta}} \tag{16.46}$$

so the problem may be restated in Fourier space, i.e., finding $Q$ such that

$$Q(\theta_1, \ldots, \theta_d) = (1-a) + \frac{1}{d} Q(\theta_1, \ldots, \theta_d) \sum_{\mu=1}^{d} \cos\theta_\mu \tag{16.47}$$

This gives, quite straightforwardly, that

$$Q(\theta_1, \ldots, \theta_d) = \frac{1-a}{1 - \frac{1}{d}\sum_{\mu=1}^{d} \cos\theta_\mu} \tag{16.48}$$

and therefore

$$q(\vec{x}) = \frac{1}{(2\pi)^d} \int_0^{2\pi} \mathrm{d}\theta_1 \ldots \int_0^{2\pi} \mathrm{d}\theta_d \frac{1-a}{1 - \frac{1}{d}\sum_{\mu=1}^{d} \cos\theta_\mu} e^{-i\vec{x}\cdot\vec{\theta}} \tag{16.49}$$

7. When starting from the origin, then of course we can reach the origin between time steps $0$ and $t$.

It remains to determine $a$. This may be done by using $q(\vec{0}) = 1$ in the above equation. This gives

$$1 = \frac{1-a}{(2\pi)^d} \int_0^{2\pi} \mathrm{d}\theta_1 \ldots \int_0^{2\pi} \mathrm{d}\theta_d \frac{}{1 - \frac{1}{d}\sum_{\mu=1}^{d} \cos\theta_\mu} \tag{16.50}$$

and therefore

$$a = 1 - \frac{1}{\mathcal{I}_d} \qquad \mathcal{I}_d = \frac{1}{(2\pi)^d} \int_0^{2\pi} \mathrm{d}\theta_1 \ldots \int_0^{2\pi} \mathrm{d}\theta_d \frac{1}{1 - \frac{1}{d}\sum_{\mu=1}^{d} \cos\theta_\mu} \tag{16.51}$$

### Critical dimension

Let us consider how this probability, and in particular $a$, changes with dimension. For $d = 1$, we get

$$\mathcal{I}_1 = \frac{1}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\theta}{1 - \cos\theta} = \infty \tag{16.52}$$

This indeed diverges for $\theta \to 0$ as the integrand scales as $\theta^{-2}$. Consequently, $a = 1$ so the walker necessarily reaches the origin when starting from any neighbouring point. It necessarily reaches this point when starting from one of its own neighbours, and so on and so forth, so that in this case $q(\vec{x}) = 1$ whatever the starting position. In dimension 2, we have the same conclusion because

$$\mathcal{I}_2 = \frac{1}{4\pi^2} \int_0^{2\pi} \mathrm{d}\theta_1 \int_0^{2\pi} \mathrm{d}\theta_2 \frac{1}{1 - \frac{1}{2}(\cos\theta_1 + \cos\theta_2)} \sim \int_{-\epsilon}^{\epsilon} \mathrm{d}\theta_1 \int_{-\epsilon}^{\epsilon} \mathrm{d}\theta_2 \frac{4}{\theta_1^2 + \theta_2^2} = \int_0^{\epsilon} \frac{\rho \mathrm{d}\rho}{\rho^2} = \infty \tag{16.53}$$

where the last transformation is a change from cartesian $(\theta_1, \theta_2)$ to polar coordinates ($\rho$ is the radius). There too, $a = 1$ and the same conclusion is reached as in dimension 1, $q(\vec{x}) = 1$. In dimension 3, the integral is convergent [8] because of the finite value of

$$\int_0^{\epsilon} \frac{\rho^2 \mathrm{d}\rho}{\rho^2} < \infty \tag{16.54}$$

Note that he average number of times the walker passes through the origin when starting from $\vec{x}$ is infinite when $d = 1$ or $d = 2$ and is finite if $d > 2$. Here again the dimension $d = 2$ marks a transition and appears as a critical dimension. Note also that in the limit of a large initial distance from the origin ($||\vec{x}|| \gg 1$) we have

$$q(\vec{x}) \sim \frac{1}{||\vec{x}||^{d-2}} \tag{16.55}$$

8. In fact, a precise computation yields $\mathcal{I}_3 \approx 1.516386$, leading to $a \approx 0.340537$. This value is less than 1 and so is $q(\vec{x})$. Note that $a > 1/6$.

## 16.4 From random walks to the diffusion equation

### 16.4.1 Generalization to continuous time

In the discrete case of a regular lattice, we had jumps from $\vec{x}$ to $\vec{x}+\vec{e}_i$ in time steps $t \to t+1$, each with probability $1/(2d)$ if unbiased. This led to Eq. (16.3), which we rewrite here, with a slightly altered notation,

$$P_{t+1}(\vec{x}) = \sum_{i=1}^{2d} \frac{1}{2d} P_t(\vec{x}+\vec{e}_i) \tag{16.56}$$

This may be generalized to a continuous time by considering a time interval $\delta t \ll 1$ and writing that there is a probability $1-\delta t$ to remain at the same position $\vec{x}$ during this time interval and a probability $\delta t/(2d)$ to jump to any neighbouring point [9]. This gives, using the notation $P_t^{\rm c}$ to specify that we are dealing with a continuous time

$$P^{\rm c}_{t+\delta t}(\vec{x}) = (1-\delta t) P^{\rm c}_t(\vec{x}) + \sum_{i=1}^{2d} \frac{\delta t}{2d} P^{\rm c}_t(\vec{x}+\vec{e}_i) \tag{16.57}$$

and we can introduce the time-derivative of $P_t^{\rm c}$ as

$$\frac{{\rm d}P^{\rm c}_t}{{\rm d}t} = \lim_{\delta t \to 0} \frac{P^{\rm c}_{t+\delta t}(\vec{x}) - P^{\rm c}_t(\vec{x})}{\delta t} = -P^{\rm c}_t(\vec{x}) + \frac{1}{2d} \sum_{i=1}^{2d} P^{\rm c}_t(\vec{x}+\vec{e}_i) \tag{16.58}$$

We can solve this by the same method as before, i.e., searching for the Fourier transform of $P_t^{\rm c}$. This gives the following form

$$P^{\rm c}_t(\vec{x}) = \frac{1}{(2\pi)^d} \int_0^{2\pi} {\rm d}\theta_1 \ldots \int_0^{2\pi} {\rm d}\theta_d \exp\left[-t\left(1 - \frac{1}{d}\sum_{\mu=1}^{d} \cos\theta_\mu\right)\right] e^{-i\vec{\theta}\cdot\vec{x}} \tag{16.59}$$

We recall that for discrete time steps we had a slightly different form, Eq. (16.17),

$$P_t(\vec{x}) = \frac{1}{(2\pi)^d} \int_0^{2\pi} {\rm d}\theta_1 \ldots \int_0^{2\pi} {\rm d}\theta_d \left(\frac{1}{d}\sum_{\mu=1}^{d} \cos\theta_\mu\right)^t e^{-i\vec{\theta}\cdot\vec{x}} \tag{16.60}$$

Let us consider a time interval $T$, and ask what is the probability that the walker has made $t$ jumps during that time. We write this probability as $Q_t(T)$. Let us split the time interval in $N = T/\delta t$ time intervals of length $\delta t$ and we shall take the limit $\delta t \to 0$. Obviously, the probability to have made no jumps is

$$Q_0 = (1-\delta t)^N = (1-\delta t)^{T/\delta t} = \exp\left[\frac{T}{\delta t}\ln(1-\delta t)\right] \approx e^{-T} \tag{16.61}$$

the last step being taken in the limit $\delta t \to 0$. More generally, we have the probability to have made $t$ jumps during time interval $T$ as

$$Q_t(T) = \frac{N!}{t!(N-t)!}(1-\delta t)^{N-t}\,\delta t^t \tag{16.62}$$

By expressing $\delta t = T/N$, we have

$$Q_t(T) = \frac{T^t}{t!} \frac{N(N-1)\ldots(N-t+1)}{N^t} \exp\left[\left(\frac{T}{\delta t} - t\right)\ln(1-\delta t)\right] \tag{16.63}$$

9. The sum of probabilities being one as it should.

In the limit $\delta t \to 0$ and $N \to \infty$, this gives

$$Q_t(T) \approx \frac{T^t}{t!} e^{-T} \tag{16.64}$$

To relate the continuous-time to the discrete-time probabilities, we note that the continuous-time probability to be at $\vec{x}$ at time $T$ is the sum of the probabilities to have reached that point in $t$ steps, summed over the number of steps. Each of these is $Q_t(T)P_t(\vec{x})$, the product of the probability to have jumped $t$ times and of the (discrete-time) probability to have reached $\vec{x}$ at the $t^{\text{th}}$ step, so that

$$P_T^{\text{c}}(\vec{x}) = \sum_t Q_t(T) P_t(\vec{x}) = \sum_t \frac{1}{(2\pi)^d} \int_0^{2\pi} \mathrm{d}\theta_1 \ldots \int_0^{2\pi} \mathrm{d}\theta_d \frac{T^t}{t!} e^{-T} \left( \frac{1}{d} \sum_{\mu=1}^{d} \cos\theta_\mu \right)^t e^{-i\vec{\theta}\cdot\vec{x}} \tag{16.65}$$

and we recognize a series expansion of the exponential as

$$\sum_t \frac{1}{t!} \left( \frac{T}{d} \sum_{\mu=1}^{d} \cos\theta_\mu \right)^t = \exp\left( \frac{T}{d} \sum_{\mu=1}^{d} \cos\theta_\mu \right) \tag{16.66}$$

This shows that we recover, as expected, the expression of $P_T^{\text{c}}$ obtained above.

### 16.4.2 Diffusion equation

The master equation on a discrete lattice for continuous time may be rewritten as

$$\frac{\mathrm{d}P_t^{\text{c}}}{\mathrm{d}t} = \frac{1}{2d} \sum_{i=1}^{2d} [P_t^{\text{c}}(\vec{x} + \vec{e}_i) - P_t^{\text{c}}(\vec{x})] \tag{16.67}$$

Now suppose that $P_t^{\text{c}}$ varies slowly with position, then a Taylor expansion yields

$$P_t^{\text{c}}(\vec{x} + \vec{e}_i) = P_t^{\text{c}}(\vec{x}) + \frac{\partial P_t^{\text{c}}}{\partial x_i} + \frac{1}{2} \frac{\partial^2 P_t^{\text{c}}}{\partial x_i^2} + \ldots \tag{16.68}$$

and the sum of two terms $P_t^{\text{c}}(\vec{x} + \vec{e}_i) - P_t^{\text{c}}(\vec{x})$ corresponding to the two opposite directions along the same coordinate leaves only the second-order derivative standing, leading to

$$\frac{\mathrm{d}P_t^{\text{c}}}{\mathrm{d}t} = \frac{1}{2d} \left( \frac{\partial^2}{\partial x_1^2} + \ldots + \frac{\partial^2}{\partial x_d^2} \right) P_t^{\text{c}}(\vec{x}) \tag{16.69}$$

where we recognize a **diffusion equation** with a diffusion coefficient $1/(2d)$.

## 16.5 Biased random walks

Among the many generalizations possible of the typical random walks we have discussed so far, the case of **biased** random walks is important. Suppose then that the probabilities to jump left or right are unequal (we suppose that $d = 1$), with a higher probability $p > 1/2$ of jumping to the right ($x$ increasing) than of jumping to the left ($q = 1 - p < 1/2$, $x$ decreasing). Then, the master equation becomes

$$P_{t+1}(x) = pP_t(x-1) + (1-p)P_t(x+1) \tag{16.70}$$

with the initial condition $P_0(x) = \delta_{x,0}$. The method of solution is identical, introducing the generating function

$$Q_t(\lambda) = \sum_x \lambda^x P_t(x) \tag{16.71}$$

for which we have, from the master equation,

$$Q_{t+1}(\lambda) = \left(p\lambda + \frac{1-p}{\lambda}\right) Q_t(\lambda) \tag{16.72}$$

and therefore, by iteration, and since $Q_0(\lambda) = 1$,

$$Q_t(\lambda) = \left(p\lambda + \frac{1-p}{\lambda}\right)^t \tag{16.73}$$

This leads to the following expression of the probability to reach $x$ at time $t$,

$$P_t(x) = C_t^{(t+x)/2} p^{(t+x)/2} (1-p)^{(t-x)/2} \tag{16.74}$$

The moments of the walker's position are then found using the same expression as before, Eq. (16.28),

$$\langle x \rangle = (2p-1)t \qquad \langle (x - \langle x \rangle)^2 \rangle = t[1 - (2p-1)^2] \tag{16.75}$$

The effect of the bias on the return to the origin is quite significant. For instance, the probability to be at the origin at time $t$ for large $t$ is

$$P_t(0) \approx \sqrt{\frac{2}{\pi t}} [4p(1-p)]^{t/2} \tag{16.76}$$

which is $1/\sqrt{t}$ in the unbiased case but goes down exponentially in time if $p \neq 1/2$. Writing $\langle n_t \rangle$ for the average number of times the walk returns to the origin during time $t$, we have

$$\langle n_t \rangle = \sum_{t_1=1}^{t} P_{t_1}(0) \tag{16.77}$$

and this goes to infinity in the limit $t \to \infty$ for an unbiased walk but goes to a finite limit if $p \neq 1/2$.