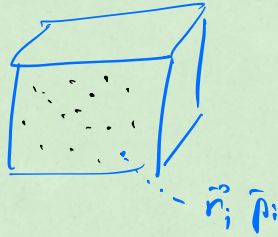


Lecture 14.11.2022

c, gibbs ensemble



V, N

gas.

6N-dim defined $r_{i,\alpha} p_{i,\alpha}$
 $i = 1 \dots N$
 $\alpha = x, y, z$

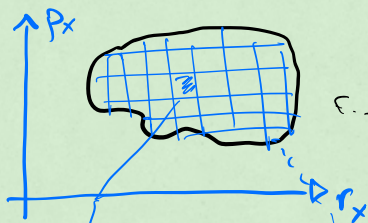
Γ -Raum

$$\int \rho(\{r_{i,\alpha}\}, \{p_{i,\alpha}\}) d^3r d^3p$$

→ # of micro states that represent a macro state ...

DENSITY FUNCTION

(p_x, p_y, p_z) (r_x, r_y, r_z) → 6-dim.



Physics.
 $E = \text{const.}$
 $V = \text{const.}$
 $N = \text{const.}$

n_i, ϵ_i

N_c # of cells in the accessible region.

2 cells

cell 1	cell 2
1 2 3	4
4 1 2	3
3 4 1	2
2 3 4	1

$N = 4$

$n_1 = 3 \quad n_2 = 1 \quad 4 \text{ ways}$

$n_1 = 2 \quad n_2 = 2 \quad 6 \text{ ways}$

$$\frac{N!}{n_1! n_2!}$$

$$n_1 + n_2 = N$$

$\{n_i\} = n_1, \dots, n_{N_c}$ set of occupation numbers.

$$\begin{aligned} \textcircled{\Delta} \quad \sum_i n_i &= N \\ \sum_i n_i \epsilon_i &= E \quad \epsilon_i = \frac{p_i^2}{2m} \end{aligned}$$

$$\Gamma_E(\{n_i\}) \propto \frac{N!}{n_1! \dots n_{N_c}!}$$

The Equilibrium is associated with the conf. $\{\bar{n}_i\}$ that maximise Γ_E Under the constraints $\textcircled{\Delta}$

density function ρ so defined that

$$\rho(\{p_{\alpha,i}\}, \{r_{\alpha,i}\}, t) d^{3N} r d^{3N} p \quad (5.1)$$

is the number of representative points that at time t are contained in the infinitesimal volume $d^{3N} r d^{3N} p$ of the Γ -Raum. Macroscopic configurations that are most likely observed shall occupy larger *volumes* in the Γ -Raum.

Microcanonical ensemble

We sketch in this section the derivation proposed by Boltzmann to estimate the *equilibrium* distribution of velocities for the molecules of an ideal gas. We consider a gas composed of N particles enclosed in a box of volume V . We assume that particles make elastic collisions among themselves and against the walls of the box, so that kinetic energy is conserved. The three components of position coordinates r_α and those of the relative momenta p_α (with $\alpha = x, y, z$) define a 6-dimensional space. For a given total energy E and volume V particles have access only to a finite portion of this space¹. Let us imagine to divide this region into N_c cells, each one of volume $\omega_c = d^3 r d^3 p$ and associate an arbitrary distribution function with the number of particles n_i that occupy the i th cell. These are called occupation numbers and need to fulfill the constraints:

$$\begin{aligned} \sum_i^{N_c} n_i &= N \\ \sum_i^{N_c} n_i \epsilon_i &= E \end{aligned} \quad (5.2)$$

where ϵ_i is the energy of the particle in the i th cell²

$$\epsilon_i = \frac{\vec{p}^2}{2m}. \quad (5.3)$$

An arbitrary set of integers n_i defines a distribution function

$$f_i = \frac{n_i}{\omega_c}. \quad (5.4)$$

This distribution is uniquely determined once we choose a point in the Γ -Raum but not vice versa. In other words, f_i is also associated with a finite

¹Technically, this corresponds to the *microcanonical ensemble*.

²Note that in the presence of a conservative force one could still define $\epsilon_i = \vec{p}^2/(2m) + \varphi(\vec{r}_i)$, with $\varphi(\vec{r}_i)$ being the potential associated with the force field.

volume of the Γ -Raum. Following Boltzmann, we denote with $\Gamma_E[\{n_i\}]$ the volume associated with a given set of occupation numbers $\{n_i\}$. This volume shall be proportional to the number of ways of distributing N distinguishable particles among the N_c cells, namely

$$\Gamma_E[\{n_i\}] \propto \frac{N!}{n_1!n_2!\dots n_{N_c}!} \quad (5.5)$$

Taking the logarithm of this expression we obtain

$$\begin{aligned} \ln \Gamma_E[\{n_i\}] &= \ln(N!) - \sum_{i=1}^{N_c} \ln(n_i!) + \text{constant} \\ &\simeq N \ln(N) - \sum_{i=1}^{N_c} n_i \ln(n_i) + \text{constant} \end{aligned} \quad (5.6)$$

where in the second row we made the approximation $\ln(n_i!) \simeq n_i \ln(n_i) - 1$ (and $\ln(N!) \simeq N \ln(N) - 1$) under the assumption that the relevant n_i be very large. The equilibrium distributions shall correspond to the set $\{n_i\}$ that maximizes the volume occupied in the Γ -Raum under the constraints given in Eqs. (5.2). Mathematically, this can be computed with the method of Lagrange multipliers, that consists in maximizing the functional

$$\Omega = N \ln(N) - \sum_{i=1}^{N_c} n_i \ln(n_i) - \lambda_\mu \left(\sum_{i=1}^{N_c} n_i - N \right) - \lambda_\beta \left(\sum_{i=1}^{N_c} n_i \epsilon_i - E \right) \quad (5.7)$$

where λ_μ and λ_β are Lagrange multipliers. The condition $\partial\Omega/\partial n_j = 0$ yields

$$-\ln(n_j) - 1 - \lambda_\mu - \lambda_\beta \epsilon_j = 0 \quad \Rightarrow \quad \bar{n}_j = C' e^{-\lambda_\beta \epsilon_j} \quad (5.8)$$

The Hessian matrix of Ω evaluated for $\{\bar{n}_j\}$ reads

$$\frac{\partial^2 \mathcal{F}}{\partial n_k \partial n_j} = -\frac{\delta_{k,j}}{\bar{n}_k}, \quad (5.9)$$

which indicates that the set $\{\bar{n}_j\}$ indeed maximizes the volume Γ_E . Setting to zero the derivative w.r.t. the Lagrange multipliers gives back the constraints in Eqs. (5.2). Formally, this means that the specific values of λ_μ and λ_β need to be adjusted to fulfill those constraints themselves. We will come back to this point later.

Perhaps the most famous Boltzmann's contribution was that of establishing a relation between the volume occupied in the Γ -Raum by a microcanonical ensemble and the relative entropy

$$S = k_B \ln \Gamma_E. \quad (5.10)$$

This equation (reproduced on Boltzmann's grave!) establishes the connection between the statistical description of microscopic states – defined within the framework of a given model – and macroscopic thermodynamics.

Thermodynamic and information-theory entropy

We make now a small digression and rewrite Eq. (5.10) in a different way. Let us replace the second row of Eq. (5.6) into the expression of the Boltzmann entropy, without imposing the constraints (5.2) and neglecting irrelevant constants

$$S = k_B \left[N \ln(N) - \sum_{i=1}^{N_c} n_i \ln(n_i) \right] = -Nk_B \sum_{i=1}^{N_c} \frac{n_i}{N} \ln \left(\frac{n_i}{N} \right). \quad (5.11)$$

- k_B N Σ: n_i/N ln(n_i/N) p_i

For a set $\{n_j\}$ the ratio n_i/N give the relative occupancy of the i th cell. If one understands n_i as random variables, those ratios represent the probability p_i that the canonical coordinates of a particle fall into the i th cell. In terms of these probabilities, the entropy per particle can then be expressed as

$$S_{s.p.} = -k_B \sum_{i=1}^{N_c} p_i \ln(p_i). \quad \text{Gibbs Entropy (5.12)}$$

(where we replaced in Eq. (5.11) n_i/N by p_i and divided by N). For a classical system with discrete set of microstates characterized by the energy E_i , p_i can be interpreted as the probability of occurrence of a given microstate. The quantity $S_{s.p.}$ is named *Gibbs entropy* and remains meaningful even when the considered system is not at thermal equilibrium.

A further generalization was (nearly unconsciously) made by C. Shannon, who defined the same quantity roughly 70 years after Boltzmann in the context of information theory

$$H[p(x)] = - \sum_x p(x) \log_2 p(x) \quad (5.13)$$

where $p(x)$ is a generic probability distribution and x a random variable that may also take continuous values (in which case the sum must be replaced with an integral). As it is not bounded to have a physical interpretation, the Shannon entropy $H[p(x)]$ lacks a proper unit. In fact, changing the base of the logarithm is equivalent to changing the unit and in information theory \log_2 is normally used because with this choice $H[p(x)]$ gives the number of bits that are necessary to store a certain piece of information. The quantity

$$I(x) = - \log_2 p(x) \quad (5.14)$$

is called information, because it measures the knowledge we gain identifying the outcome of a random trial. For instance, by tossing a coin and getting a head outcome, we gain an amount of information equal to $\log_2(2)$; by casting a die and getting 3 as an outcome we gain an amount of information equal to $\log_2(6)$, etc. The Shannon entropy is the expected information averaged over all the possible events and is a property of the probability distribution $p(x)$ itself.

The above derivation clearly shows the analogy between information-theory entropy and thermodynamic entropy³. In this regard, it is worth quoting what Shannon said about the entropy named after him:

My greatest concern was what to call it (...). When I discussed it with John von Neumann (...) he told me, "You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one knows what entropy really is, so in a debate you will always have the advantage."

But can we learn something about thermodynamic entropy from Shannon entropy? In the attempt to get little insight into this question, suppose we are dealing with a system of 3 spins $1/2$. The relative Hilbert space consists of $2^3 = 8$ states that we associate with the microrstates of the system. Putting aside any physical meaning or energetic consideration, we may map those states into 8 different cards. The probability to pick a card i at random shall be $p_i = 1/8$, which yields the Shannon entropy

$$H[p(x)] = - \sum_x^8 \frac{1}{8} \log_2 \left(\frac{1}{8} \right) = \log_2 8 = 3. \quad (5.15)$$

This reflects the fact that, if someone choose at random one card out of 8, I need to pose *at least* 3 questions in order to identify the chosen card without having any prior knowledge (or using any trick). For this trivial example, we see that the Shannon entropy of the cards problem is equivalent to the Gibbs entropy of a system of 3 non-interacting spins $1/2$, with all degenerate eigenvalues and at thermodynamic equilibrium (provided we change the unit with which entropy is measured by a factor $k_B/\log_2 e$). Let us now assume that our 3 spins are coupled through a ferromagnetic exchange interaction of Heisenberg type; we know that the ground state is realized by the multiplet with maximal spin $S^T = 3/2$. Then, if the strength of the exchange interaction is much larger than the thermal energy we can assume that only the ground state be populated. This restricts the accessible states to 4 and

³See the manuscript [CourseLibrary/Articles/Aikaike.Prediction_and_Entropy.pdf](#) for a historical account.

No info about system all 8 states are equally likely

3 spin 1/2

dim Hilbert $2^3 = 8$

CARDS

joke

$p_i = \frac{1}{8}$ 8 states equally likely

$$H(p_i) = - \sum_{i=1}^8 p_i \log_2 p_i$$

$$= - \sum \frac{1}{8} \log_2 \left(\frac{1}{8} \right)$$

$$= - \frac{1}{8} \log_2(8) =$$

$$- \log_2(8) = 3$$

No Physics

Q1, Q2, Q3
I need to ask 3 questions with no other info

S^T	k	M	
$3/2$		$+3/2$	4
		$+1/2$	2
		$-1/2$	1
		$-3/2$	3
$1/2$	$2\pi/3$	$+1/2$	2
 Q3		$-1/2$	1
$1/2$	$-2\pi/3$	$+1/2$	2
		$-1/2$	1

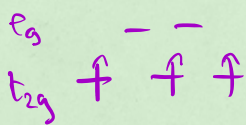
Q1

not 5/2

Q2

Physical Information

Let's say 3 spins are the electron of Cr^{3+} in Oh



Assume intraatomic coupling can be described $\mathcal{H} = -J_{intra} (\hat{S}_1 \hat{S}_2 + \hat{S}_2 \hat{S}_3 + \hat{S}_3 \hat{S}_1)$

we know that $J_{intra} \gg k_B T$

\Rightarrow only $S^T = 3/2$ multiplet is populated
only 4 states with the same energy

$$p_i = \frac{1}{4} \text{ and } i=1,2,3,4 \quad H(p_i) = - \sum p_i \log_2 \left(\frac{1}{4} \right) =$$

TAKE gaining (ferromagnetic) info about the system Shannon entropy is reduced!

If in addition I apply a field $\mu_B B \gg k_B T$

\Rightarrow only state with $|3/2, -3/2\rangle$ is populated

$$\Rightarrow p_1 = 1 \quad \text{and} \quad p_i = 0 \quad \forall i \neq 1$$

$$\Rightarrow H\{p_i\} = \dots \log_2(1) = 0 \quad \text{Entropy vanishes!}$$

CHAPTER 5. THERMODYNAMICS OF NON-INTERACTION PARTICLES 77

accordingly the Shannon entropy associated with those states reduces to

$$H[p_i] = - \sum_x \frac{1}{4} \log_2 \left(\frac{1}{4} \right) = \log_2 4 = 2. \quad (5.16)$$

If we then apply a magnetic field and hypothetically achieve a Zeeman splitting much larger than the thermal energy ($\mu_B B \gg k_B T$), only one state of the multiplet $S^T = 3/2$ will be significantly populated: the corresponding entropy is thus suppressed to zero. From this conceptual experiment we learn that introducing physical knowledge about a system reduces the Shannon entropy or, equivalently, increases our knowledge about the microstates of the system. In this perspective, we may regard the entropy of a system as the amount of “missing” information needed to determine a microscopic state from the knowledge of its state at the level of macroscopic thermodynamics⁴

In the quantum-mechanical description the Gibbs entropy is replaced by the von Neumann entropy defined in terms of the density matrix.

Canonical ensemble

Following the textbook by K. Huang, we now ask ourselves what ensemble is appropriate to describe a system not isolated but in thermal equilibrium with a larger one. The goal is to determine the probability of finding this system in a given point in the Γ -Raum, namely to define the functional dependence of the density function $\rho(\{p_{\alpha,i}\}, \{r_{\alpha,i}\})$ (with $i \in$ system 1). To this purpose, we consider a system described by the microcanonical ensemble with energy E and split it into two subsystems 1 and 2. These two subsystems can exchange energy but we shall assume that their average energies \bar{E}_1 and \bar{E}_2 be defined and $\bar{E}_1 \ll \bar{E}_2$. The probability of the system 1 being represented by a point in the elementary volume $d\Gamma_1$ of the Γ -Raum characterized by the energy E_1 is proportional to the volume occupied by all the microstates of the system 2 associated with the energy $E_2 = E - E_1$ (by definition of E). In formula, we express this fact as

$$\rho(\{p_{\alpha,i}\}, \{r_{\alpha,i}\}) \propto \Gamma_2(E_2) = \Gamma_2(E - E_1). \quad (5.17)$$

Equation (5.35) relates $\Gamma_2(E - E_1)$ to the entropy $S_2(E - E_1)$ associated with the system 2. Since we expect only values $E_1 \simeq \bar{E}_1$ to be important and we

⁴The interested reader is encouraged to read the article by E. T. Jaynes, Phys. Rev. **106** 620 (1957) available in the *CourseLibrary*, in which a reversed viewpoint is taken: entropy and its maximization are assumed as the starting concept; the usual results of statistical mechanics are then derived imposing different physical constraints case by case, by means of the Lagrange multipliers approach.

have assumed that $\bar{E}_1 \ll \bar{E}_2$, we can expand the entropy about the value $E = E_2$ to obtain the relation

$$k_B \ln [\Gamma_2(E - E_1)] = S_2(E - E_1) \simeq S_2(E) - E_1 \left[\frac{\partial S_2}{\partial E} \right]_{E=E_2}. \quad (5.18)$$

The derivative of the entropy w.r.t. to the energy is the definition of the inverse temperature in the microcanonical ensemble; in the specific case it is the inverse temperature $1/T_2$ of the largest system. Therefore, neglecting the first term independent of E_1 and T_2 the equation above implies that

$$\Gamma_2(E - E_1) \propto \exp \left[-\frac{E_1}{k_B T_2} \right]. \quad (5.19)$$

Since we have assumed that the systems are in thermal equilibrium, it is $T_2 = T_1 = T$. Moreover, we know that for a classical system $E_1 = \mathcal{H}(\{p_{\alpha,i}\}, \{r_{\alpha,i}\})$, thus the sought for density function ρ is

$$\rho(\{p_{\alpha,i}\}, \{r_{\alpha,i}\}) = e^{-\beta \mathcal{H}(\{p_{\alpha,i}\}, \{r_{\alpha,i}\})} \quad (5.20)$$

with $\beta = 1/(k_B T)$. Note that the temperature T (contained in β) summarizes all the information about the degrees of freedom of the larger system 2 that acts as a heat reservoir. The ensemble defined by the density function in Eq. (5.20) is called *canonical ensemble*. The energy levels of a system at thermal equilibrium should fulfill the Boltzmann statistics expressed by this density function. In the following, we will actually use this as definition of *thermal equilibrium*.

Thermal averages

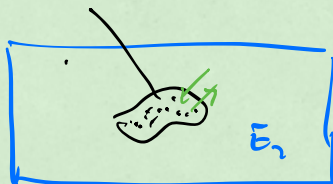
In the remainder of the course we will largely use the Boltzmann distribution (5.20) to compute thermal averages in the canonical ensemble. In this approach a fundamental quantity is the partition function, which for a gas reads

$$\mathcal{Z} = \frac{1}{N!} \int \frac{d^{3N} r d^{3N} p}{(2\pi\hbar)^{3N}} e^{-\beta \mathcal{H}(\{p_{\alpha,i}\}, \{r_{\alpha,i}\})}, \quad (5.21)$$

where the factor $N!$ accounts for the “correct Boltzmann counting”. This correction is necessary only when a system of particles that are not distinguishable at a macroscopic level is considered. \mathcal{Z} is related to a thermodynamic potential, \mathcal{F} , via the general relation

$$\mathcal{F} = -\frac{1}{\beta} \ln \mathcal{Z}. \quad (5.22)$$

E_1 not constant



$E_2 > E_1$

$T_1 = T_2 = T$

Canonical ensemble

The average of any observable $\mathcal{O}(\{p_{\alpha,i}\}, \{r_{\alpha,i}\})$ can be computed as

$$\langle \mathcal{O} \rangle = \frac{1}{\mathcal{Z}} \frac{1}{N!} \int \frac{d^{3N}q d^{3N}p}{(2\pi\hbar)^{3N}} \mathcal{O}(\{p_{\alpha,i}\}, \{r_{\alpha,i}\}) e^{-\beta\mathcal{H}(\{p_{\alpha,i}\}, \{r_{\alpha,i}\})}. \quad (5.23)$$

A useful trick is that of adding to the Hamiltonian a term in which a (conjugated) field ϕ is coupled with the observable one would like to compute:

$$\mathcal{H}(\{p_{\alpha,i}\}, \{r_{\alpha,i}\}) \Rightarrow \mathcal{H}(\{p_{\alpha,i}\}, \{r_{\alpha,i}\}) - \phi \mathcal{O}(\{p_{\alpha,i}\}, \{r_{\alpha,i}\}). \quad (5.24)$$

Given the equivalence

$$\left[\frac{\partial e^{-\beta(\mathcal{H}+\phi\mathcal{O})}}{\partial\phi} \right]_{\phi=0} = -\beta\mathcal{O}e^{-\beta\mathcal{H}} \quad (5.25)$$

one has that

$$\langle \mathcal{O} \rangle = -\frac{1}{\beta} \frac{1}{\mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial \phi} = -\frac{1}{\beta} \frac{\partial \ln \mathcal{Z}}{\partial \phi}. \quad (5.26)$$

If one is interested in the thermal average of the Hamiltonian itself, the relevant relation is

$$\mathcal{H} = -\frac{\partial e^{-\beta\mathcal{H}}}{\partial\beta} \quad (5.27)$$

and thus

$$\langle \mathcal{H} \rangle = -\frac{\partial \ln \mathcal{Z}}{\partial\beta}. \quad (5.28)$$

The thermodynamic potential \mathcal{F} will inherit from the partition function \mathcal{Z} the dependence on external parameters. For an ideal gas these parameters are (N, V, T) , corresponding to potential called Helmholtz free energy. For a magnetic system the typical external parameters are (N, B, T) and the associated potential is called Gibbs free energy, usually indicated with G . To avoid proliferation of notation, hereafter we shall use the symbol \mathcal{F} and call it just *free energy*, implicitly assuming that it represents the thermodynamic potential dependent on the parameters of our model.

Further details about how to compute thermal averages in the canonical ensemble are given in the Appendix.

Equipartition theorem for an ideal gas

For an ideal gas the partition function given in Eq. (5.21) reduces to the product of N independent, single-particle partition functions:

$$\mathcal{Z} = \frac{1}{N!} \mathcal{Z}_{\text{s.p.}}^N \quad (5.29)$$

$$f(\{p_i\}, \{r_i\}) = e^{-\beta H} \quad \text{Boltzmann}$$

$$Z = \left(\frac{1}{N!} \right) \int \frac{d^{3N}p d^{3N}r}{(2\pi\hbar)^{3N}} e^{-\beta H} \quad \text{Partition function}$$

Correct Boltzmann counting

$$\langle \Phi \rangle = \frac{1}{Z} \frac{1}{N!} \int \frac{d^{3N}p d^{3N}r}{(2\pi\hbar)^{3N}} \Phi e^{-\beta H}$$

$$\langle H \rangle = - \frac{1}{Z} \frac{\partial Z}{\partial \beta} = - \frac{\partial \ln Z}{\partial \beta}$$

$H \mapsto H - \Phi$ field conjugated to the observable Φ

$$\frac{\partial}{\partial \Phi} e^{-\beta(H + \Phi)} = -\beta \Phi e^{-\beta(H + \Phi)}$$

$$\langle \Phi \rangle = - \beta \frac{1}{Z} \frac{\partial Z}{\partial \Phi} = - \beta \ln(Z)$$

Boltzmann microcanonical.

$$\bar{n}_i \propto e^{-\lambda \beta \epsilon_i}$$

$\lambda \beta$ adjusted to

$$\sum_i \bar{n}_i \epsilon_i = E$$

Canonical ensemble

$$\langle H \rangle = - \frac{\partial}{\partial \beta} \ln Z$$

$$I = \int_{\mathbb{R}} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}}$$

$$-\frac{\partial}{\partial \alpha} I = \int_{\mathbb{R}} x^2 e^{-\alpha x^2} dx = \frac{1}{2} \frac{\sqrt{\pi}}{\alpha^{3/2}}$$

$$H = \sum_{i,\alpha} \frac{p_{i,\alpha}^2}{2m} \Rightarrow \langle H \rangle = \sum_{i,\alpha} \left\langle \frac{p_{i,\alpha}^2}{2m} \right\rangle$$

$$H = \frac{1}{2} k \sum_i (x_i - x_{i+1})^2 \Rightarrow H(q) = \sum_q \left\langle \frac{\Gamma(q)}{2} |x_q|^2 \right\rangle$$

$\frac{1}{2} k_B T$
 $\frac{1}{2} k_B T$

$$Z = \frac{1}{N!} (Z_{\text{s.p.}})^N$$

CHAPTER 5. THERMODYNAMICS OF NON-INTERACTION PARTICLES 80

with

$$Z_{\text{s.p.}} = \frac{V}{(2\pi\hbar)^3} \int \Pi_{\alpha} \exp \left[-\frac{\beta p_{\alpha}^2}{2m} \right] dp_{\alpha}. \quad (5.30)$$

and $\alpha = x, y, z$. The thermal average of each quadratic term in the Hamiltonian can be computed using Eq. (5.28) and the well-known result

$$\int du_{\alpha} \exp \left[-\frac{\beta}{2m} (u_{\alpha})^2 \right] = \sqrt{\frac{2\pi m}{\beta}}. \quad (5.31)$$

One finally obtains

$$\left\langle \frac{1}{2m} p_{\alpha}^2 \right\rangle = \frac{1}{2} k_B T \quad (5.32)$$

which is the celebrated equipartition theorem. The beauty of this theorem is that it applies to every Hamiltonian that can be decoupled into the sum of N independent quadratic degrees of freedom: each one of such degrees of freedom contributes to the total averaged energy with a term $k_B T/2$. In some future chapter, we will use this powerful result to prove that cooperative models of magnetism with continuous symmetry cannot sustain long-range order on lattices of dimension 1 or 2. For the specific case of an ideal gas composed of N particles, the total quadratic degrees of freedom – associated with the kinetic energy – are $3N$ and thus the average of the total energy reads

$$\langle \mathcal{H} \rangle = \frac{3N}{2} k_B T. \quad (5.33)$$

Microcanonical vs canonical ensemble

While deriving the equilibrium distribution of occupation numbers $\{\bar{n}_j\}$ in the microcanonical ensemble we found that

$$\bar{n}_j = C' e^{-\lambda_{\beta} \epsilon_j}. \quad (5.34)$$

Inserting Eqs. (5.3) and (5.4) into the equation above, one obtains the following distribution

$$f_i = C e^{-\lambda_{\beta} \frac{p_i^2}{2m}}. \quad (5.35)$$

In that context we wrote that the Lagrange multiplier λ_{β} is to be adjusted to fulfill the constraint that the total energy E takes a predefined value. If we assume that the average energy per particle in the microcanonical ensemble must equal the expected thermal average $3k_B T/2$, the descriptions given in the microcanonical and in the canonical ensemble must be equivalent. This requirement leads us to identify the Lagrange multiplier λ_{β} with

$\beta = 1/(k_B T)$. Note that with this choice the probability distribution f_i corresponds to the Maxwell-Boltzmann distribution of velocities.

Taking a second look at the role played by ϕ and β in Eqs. (5.26) and (5.28), the reader will note that the Lagrange multipliers introduced in the microcanonical ensemble to impose constraints are mapped into conjugated fields used to compute thermal averages in the canonical or in the grand canonical ensemble⁵

As a small concluding remark we note that to derive the density function in Eq. (5.20) we assumed the coupling between the two parts 1 and 2 of the whole system to be small. This implicitly relies on the interaction between particles to be short-ranged enough. For instance, for a system of charged particles interacting via Coulomb interaction this is not true and the physics described in the canonical ensemble may differ from that given in the microcanonical ensemble: The latter is in such cases the only meaningful framework to describe the whole system.

5.2 Non-interacting magnetic moments

The Bohr–van Leeuwen theorem

We turn now our attention to the problem of $N \sim 10^{23}$ non-interacting electrons in a solid experiencing a constant magnetic field. We will treat them as a gas of classical particles and associate with each electron a single-particle Hamiltonian of the form

$$\mathcal{H} = \frac{1}{2m} \left(\vec{p} - q_e \vec{A} \right)^2 + q_e \phi(r) \quad (5.36)$$

with $q_e = -e < 0$. The electrostatic potential $\phi(r)$ is generated by the nuclei and other charges in the crystal, but it is assumed to be independent of the reciprocal position of electrons⁶. We would like to compute the thermal average of the magnetic moment per electron in response to the application of an external field or, in other words, the average polarization of the electron gas at finite temperature. For classical particles the partition function in the canonical ensemble is given by Eq. (5.21). The average magnetic moment

⁵The role of the Lagrange multiplier λ_μ introduced in the microcanonical ensemble is taken by the chemical potential in the grand canonical ensemble, where it is used to compute the average number of particles in the system.

⁶Note that a realistic Hamiltonian would depend on the specific position of all the other electrons in the solid, while $\phi(r)$ is assumed to depend only on the position of the considered electron.

along the applied field is obtained as the derivative of the free energy w.r.t. to the applied field

$$\langle \mu^\alpha \rangle = -\frac{1}{N} \frac{\partial \mathcal{F}}{\partial B^\alpha} = \frac{k_B T}{N} \frac{1}{\mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial B^\alpha}, \quad (5.37)$$

with $\alpha = x, y, z$. As done for the ideal gas, we can limit ourselves to considering the partition function produced by a single particle:

$$\mathcal{Z} = \frac{1}{(2\pi\hbar)^3} \int d^3p \exp \left[-\frac{\beta}{2m} (\vec{p} - q_e \vec{A})^2 \right] \int d^3r \exp [-\beta q_e \phi(r)]. \quad (5.38)$$

The integration with respect to d^3p can be transformed into three standard Gaussian integrals with a shift of the origin

$$\begin{cases} u_x = p_x - q_e A_x \\ u_y = p_y - q_e A_y \\ u_z = p_z - q_e A_z \end{cases} \quad (5.39)$$

After this transformation of coordinates (see Appendix 5.2 for further details) one obtains

$$\mathcal{Z} = \frac{1}{\hbar^3} \left(\frac{m}{2\pi\beta} \right)^{3/2} \int d^3r \exp [-\beta q_e \phi(r)] \quad (5.40)$$

(where we have made use of the result (5.31)). Since the integral with respect to d^3r – so-called *the configuration integral* – does not depend on the magnetic field \vec{B} , the whole partition function is also independent of B . In line with Eq. (5.37), this means that the average magnetic moment (i.e. polarization of the electron gas) vanishes. In other words, if statistical mechanics and classical mechanics are applied consistently, the average magnetic moment per electron is always zero. This important result of statistical physics is named the Bohr–van Leeuwen theorem, after Niels Bohr and Hendrika Johanna van Leeuwen who proved it independently in the 1910s. The importance of this theorem was acknowledged only considerably later by van Vleck (1932):

“At any finite temperature, and in all finite applied electric or magnetic fields, the net magnetization of a collection of electrons in thermal equilibrium vanishes identically”.

Practically, this theorem marks a conclusive statement about the need for quantum mechanics to account for magnetic phenomena, like diamagnetism, paramagnetism or ferromagnetism.

constant
 \vec{E}, \vec{B}

e.s. potential
 $\vec{E} = -\vec{\nabla}\phi$
 $\vec{B} = \vec{\nabla} \times \vec{A}$