Introduction to Quantum Statistical Mechanics

Alain Joye

Institut Fourier, Université de Grenoble 1,
BP 74, 38402 Saint-Martin d'Hères Cedex, France
`alain.joye@ujf-grenoble.fr`

This set of lectures is intended to provide a flavor of the physical ideas underlying some of the concepts of Quantum Statistical Mechanics that will be studied in this school devoted to Open Qantum Systems. Although it is quite possible to start with the mathematical definitions of notions such as "bosons", "states", "Gibbs prescription" or "entropy" for example and prove theorems about them, we believe it can be useful to have in mind some of the heuristics that lead to their precise definitions in order to develop some intuition about their properties.

Given the width and depth of the topic, we shall only be able to give a very partial account of some of the key notions of Quantum Statistical Mechanics. Moreover, we do not intend to provide proofs of the statements we make about them, nor even to be very precise about the conditions under which these statements hold. The mathematics concerning these notions will come later. We only aim at giving plausibility arguments, borrowed from physical considerations or based on the analysis of simple cases, in order to give substance to the dry definitions.

Our only hope is that the mathematically oriented reader will benefit somehow from this informal introduction, and that, at worse, he will not be too confused by the many admittedly hand waving arguments provided.

Some of the many general references regarding an aspect or the other of these lectures are provided at the end of these notes.

1 Quantum Mechanics

We provide in this section an introduction to the quantum description of a physical system, starting from the Hamiltonian description of Classical Mechanics. The quantization procedure is illustrated for the standard kinetic plus potential Hamiltonian by means of the usual recipe. A set of postulates underlying the quantum description of systems is introduced and motivated by means of that special though important case. These aspects, and much more, are treated in particular in [GJ] and [MR], for instance.

1.1 Classical Mechanics

Let us recall the Hamiltonian version of Classical Mechanics in the following typical setting, neglecting the geometrical content of the formalism. Consider N particles in $\mathbb{R}^d$ of coordinates $q_k \in \mathbb{R}^d$, masses m_k and momenta $p_k \in \mathbb{R}^d$, $k = 1, \cdots, N$, interacting by means of a potential

$$\begin{aligned} V : \quad & \mathbb{R}^{dN} \to \mathbb{R} \\ & q \mapsto V(q). \end{aligned} \tag{1}$$

The space $\mathbb{R}^{dN}$ of the coordinates $(q_1, q_2, \cdots, q_N)$, with $q_{k,j} \in \mathbb{R}$, $j = 1 \cdots, d$ which we shall sometimes denote collectively by q (and similarly for p), is called the configuration space and the space $\Gamma = \mathbb{R}^{dN} \times \mathbb{R}^{dN} = \mathbb{R}^{2dN}$ of the variables (q, p) is called the phase space. A point (q, p) in phase space characterizes the *state* of the system and the *observables* of the systems, which are the physical quantities one can measure on the system, are given by functions defined on the phase space. For example, the potential is an observable. The Hamiltonian $H : \Gamma \to \mathbb{R}$ of the above system is defined by the observable

$$H(p, q) = \sum_{k=1}^{N} \frac{p_k^2}{2m_k} + V(q_1, q_2, \cdots, q_N), \tag{2}$$

$$\text{where } V(q_1, q_2, \cdots, q_N) = \sum_{i<j} V_{ij}(|q_i - q_j|),$$

which coincides with the sum of the kinetic and potential energies. The equations of motion read for all $k = 1, \cdots, N$ as

$$\dot{q}_k = \frac{\partial}{\partial p_k} H(q, p), \quad \dot{p}_k = -\frac{\partial}{\partial q_k} H(q, p), \quad \text{with} \quad (q(0), p(0)) = (q^0, p^0), \tag{3}$$

where $\frac{\partial}{\partial q_k}$ denotes the gradient with respect to q_k. The equations (3) are equivalent to Newton's equations, with $p_k = m_k \dot{q}_k$,

$$m_k \ddot{q}_k = -\frac{\partial}{\partial q_k} V(q) \quad \text{with} \quad (q(0), \dot{q}(0)) = (q^0, \dot{q}^0),$$

for all $k = 1, \cdots, N$. In case the Hamiltonian is time independent, i.e. if the potential V is time independent, the total energy E of the system is conserved

$$E = H(q(0), p(0)) \equiv H(q(t), p(t)), \quad \forall t. \tag{4}$$

where $(q(t), p(t))$ are solutions to (3) with initial conditions $(q(0), p(0))$. More generally, a system is said to be Hamiltonian if its equations of motions read as (3) above. We shall essentially only deal with systems governed by Hamiltonians that are time-independent.

These evolution equations are also called *canonical equations of motion*. Changes of coordinates

$$p_k \mapsto P_k, \quad q_k \mapsto Q_k, \text{ such that } H(q,p) \mapsto K(Q,P)$$

which conserve the form of the equations of motions, i.e.

$$\dot{Q}_k = \frac{\partial}{\partial P_k} K(Q,P), \quad \dot{P}_k = -\frac{\partial}{\partial Q_k} K(Q,P), \quad \text{with } (Q(0), P(0)) = (Q^0, P^0), \tag{5}$$

are called *canonical transformations*. The energy conservation property (4) is just a particular case of time dependence of a particular observable. Assuming the Hamiltonian is time independent, but not necessarily given by (2), the time evolution of any (smooth) observable $B : \Gamma \to \mathbb{R}$ defined on phase space computed along a classical trajectory $B_t(q,p) \equiv B(q(t), p(t))$ is governed by the equation

$$\frac{d}{dt} B_t(q,p) = L_H B_t(q,p), \quad \text{with } B_0(q,p) = B(q^0, p^0), \tag{6}$$

where the linear operator L_H acting on the vector space of observables is given by

$$L_H = \nabla_p H(q,p) \cdot \nabla_q - \nabla_q H(q,p) \cdot \nabla_p, \tag{7}$$

with the obvious notation. Observables which are constant along the trajectories of the system are called *constant of the motions*. Introducing the Lebesgue measure on $\Gamma = \mathbb{R}^{2dN}$,

$$d\mu = \Pi_{k=1}^{N} dq_k dp_k, \quad \text{with } dq_k = \Pi_{j=1}^{d} q_{k,j}, \text{ and } dp_k = \Pi_{j=1}^{d} p_{k,j},$$

and the Hilbert space $L^2(\Gamma, d\mu)$, one checks that L_H is formally anti self-adjoint on $L^2(\Gamma, d\mu)$, (i.e. antisymmetrical on the set of observables in $C_0^\infty(\Gamma)$). Therefore, the formal solution to (6) given by

$$B_t(q,p) = e^{tL_H} B_0(q,p)$$

is such that e^{tL_H} is unitary on $L^2(\Gamma, d\mu)$. Another expression of this fact is *Liouville's Theorem* stating that

$$\frac{\partial(q(t), p(t))}{\partial(q^0, p^0)} \equiv 1,$$

where the LHS above stands for the Jacobian of the transformation $(q^0, p^0) \mapsto (q(t), p(t))$. It is convenient for the quantization procedure to follow to introduce the *Poisson bracket* of observables B, C on $L^2(\Gamma)$ by the definition

$$\{B, C\} = \nabla_q B \cdot \nabla_p C - \nabla_p B \cdot \nabla_q C. \tag{8}$$

For example,

$$\{q_{k,m}, p_{j,n}\} = \delta_{(j,n),(k,m)}, \ \{p_{j,n}, p_{k,m}\} = \{q_{j,n}, q_{k,m}\} = 0, \qquad (9)$$

which are particular cases of

$$\{q_{k,m}, F(q,p)\} = \frac{\partial F(q,p)}{\partial p_{k,m}}, \ \ \{G(q,p), p_{j,n}\} = \frac{\partial G(q,p)}{\partial q_{j,n}}.$$

These brackets fulfill Jacobi's relations,

$$\{A, \{B, C\}\} + \text{circular permutations} = 0 \qquad (10)$$

and, as $\{H, B\} = -L_H B$, we can rewrite (6) by means of Poisson brackets as

$$\frac{d}{dt} B_t = -\{H, B_t\}. \qquad (11)$$

Therefore, it follows that the Poisson bracket of two constants of the motion is a constant of the motion, though not necessarily independent from the previous ones.

Before we proceed to the quantization procedure, let us introduce another Hamiltonian system we will be interested in later on. It concerns the evolution of N identical particles of mass m and charge e in $\mathbb{R}^3$, interacting with each other and with an external electromagnetic field (E, B).

Let us recall Maxwell's equations for the electromagnetic field

$$\nabla B = 0, \quad \nabla \wedge E = -\frac{\partial B}{\partial t}, \quad \epsilon_0 \nabla E = \rho_e, \quad \nabla \wedge B = \mu_0 j + \frac{1}{c^2} \frac{\partial E}{\partial t}, \quad (12)$$

where ρ_e and j denote the density of charges and of current, respectively, the constant ϵ_0 and μ_0 are characteristics of the vacuum in which the fields propagate, and c is the speed of light. A particle of mass m and charge e in presence of an electromagnetic field obeys the Newtonian equation of motion determined by the Lorentz force

$$m\ddot{q} = e(E + \dot{q} \wedge B). \qquad (13)$$

When N charged particles interact with the electromagnetic field, the rule is that each of them becomes a source for the fields and obeys (13), with the densities given by

$$\rho_e(x, t) = \sum_{j=1}^{N} e\delta(x - q_j(t)), \ \text{ and } \ j = \sum_{j=1}^{N} e\dot{q}_j(t)\delta(x - q_j(t)). \qquad (14)$$

In order to have a Hamiltonian description of this dynamics later, we introduce the scalar potential V and the *vector potential A* associated to the electromagnetic field (E, B). They are defined so that

$$E = -\frac{\partial A}{\partial t} - \nabla V, \quad B = \nabla \wedge A, \qquad (15)$$

hence the first two equations (12) are satisfied, and

$$\frac{\partial}{\partial t}(\nabla A) + \Delta V = -\epsilon_0^{-1}\rho_e$$

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}A - \Delta A + \nabla(\nabla A + \frac{1}{c^2}\frac{\partial V}{\partial t}) = \mu_0 j \qquad (16)$$

yield the last two equations of (12). There is some freedom in the choice of A and V in the sense that the physical fields E and B are unaffected by a change of the type

$$V \mapsto V + \frac{\partial \chi}{\partial t}, \quad A \mapsto A - \nabla\chi,$$

where χ is any scalar function of (x, t). A transformation of this kind is called a *gauge transform*. This allows in particular to choose the potential vector A so that it satisfies

$$\nabla A = 0, \qquad (17)$$

by picking a χ solution to $\Delta\chi = \nabla A$, if (17) is not satisfied. This is the so called *Coulomb gauge* in which (16) reduces to

$$\Delta V = -\epsilon_0^{-1}\rho_e \iff V(x,t) = \frac{1}{4\pi\epsilon_0}\int dy\frac{\rho_e(y,t)}{|x-y|}$$

$$\frac{1}{c^2}\frac{\partial^2}{\partial t^2}A - \Delta A = \mu_0 j - \frac{1}{c^2}\frac{\partial \nabla V}{\partial t} \equiv \mu_0 j_T. \qquad (18)$$

The subscript T stands here for transverse, since $\nabla j_T \equiv 0$. Assume we have N particles of identical masses m and identical charges e interacting with the electromagnetic field satisfying (18) with (14) so that

$$V(x,t) = \frac{e}{4\pi\epsilon_0}\sum_{j=1}^{N}dy\frac{1}{|x-q_j(t)|}. \qquad (19)$$

We want to write down a Hamiltonian function which will yield (13) back when we compute the equation of motions for the particles as (3). It is just a matter of computation to show that the following (time-dependent) Hamiltonian fulfills this requirement:

$$H(q,p,t) = \sum_j \frac{1}{2m}|p_j - eA(q,t)|^2 + \frac{1}{8\pi\epsilon_0}\sum_{i\neq j}\frac{e^2}{|q_i - q_j|}. \qquad (20)$$

The only thing to note is that when one computes the part of the electric field E that is produced by the ∇V part of (15) at the point $q_j(t)$, the double sum that appears due to the form (19) of V contains an infinite part when the indices take the same value, which is simply ignored .

Remark: It follows also from the above that the momentum p_j of the j'th particle isn't proportional to its velocity, but is given instead by

$$p_j = mv_j + eA(q_j, t).$$

This is an important feature of systems interacting with electromagnetic fields. As noted above, this Hamiltonian is time dependent. We can get a time independent Hamiltonian provided one takes also into account the energy of the field in the Hamiltonian. This new Hamiltonian H^{tot} reads

$$H^{tot} = H(q, p, t) + \frac{1}{2} \int dx \left(\epsilon_0 |\partial A(x, t)/\partial t|^2 + \mu_0^{-1} |\nabla \wedge A(x, t)|^2 \right),$$

where the first term in the integral is the contribution to the electric field that is not provided by the Coulomb potential (19) (which is taken into account in (20)) and the second term is the magnetic energy. It can be shown also that the total energy H^{tot} is conserved.

1.2 Quantization

The Quantum description of a general classical system is given by a set of postulates we list here as **P1** to **P4**. In order to motivate and/or illustrate their meaning, we consider in parallel the typical Hamiltonian (2) to make the link with its quantization by means of the traditional recipe.

P1: *The phase space Γ is replaced by a Hilbert space $\mathcal{H}$ whose scalar product is denoted by $\langle \cdot | \cdot \rangle$, with anti-linearity on the left. The state of the system is characterized by a ray in this space, that is a unit vectors with an arbitrary phase.*

Actually, rays characterize the *pure states* of the system. When we consider Quantum Statistical Mechanics, we will make a distinction between pure states and mixed states that will be introduced then. However, in that section, we will go on talking about states.

In case of our example, $\mathcal{H} = L^2(\mathbb{R}^{dN})$, $\mathbb{R}^{dN}$ being the configuration space. The state of the system is characterized by a normalized complex valued function $\psi(q)$ in $L^2(\mathbb{R}^{dN})$, also called the *wave function* of the system.

P2: *The observables are given by (possibly unbounded) self-adjoint linear operators on $\mathcal{H}$ obtained from their classical counterparts by a quantization procedure.*

The quantization procedure is not always straightforward. In particular, if the phase space has a non trivial topological structure, sophisticated methods of quantizations have to be applied. The link with the corresponding classical

observables should be achieved, at a formal level at least, by taking the limit $\hbar \to 0$.

For our example, the formal substitutions

$$p_k \mapsto -i\hbar\nabla_k, \quad q_k \mapsto \mathrm{mult}q_k, \quad k = 1, \cdots, N \qquad (21)$$

are used. Here $\mathrm{mult}q_k$ denotes the operator multiplication by the variable q_k, which we shall simply denote by q_k below, and $\hbar$ is Planck's constant, whose numerical value is about $1.055 \times 10^{-34} J.s$. In particular, the classical Hamiltonian (2) gives rise to the (formally) self-adjoint operator

$$H = \sum_{k=1}^{N} -\frac{\hbar^2}{2m_k}\Delta_k + V(q_1, q_2, \cdots, q_N) \text{ on } \mathcal{H}, \qquad (22)$$

where Δ_k denotes the Laplacian in the variables q_k. This class of operators goes under the name *Schrödinger operators* and plays, for obvious reasons, an important role in Quantum Mechanics. Note that the quantization of observables by the formal rule (21) may need to be precised by a symmetrization procedure due to the non-commutativity of p and q,

$$[p_{j,n}, q_{k,m}] = \frac{\hbar}{i}[\partial_{q_{j,n}}, q_{k,m}] = \frac{\hbar}{i}I\!I, \qquad (23)$$

where $I\!I$ denotes the identity operator. The symmetrization can be performed by hand in some concrete cases. For example, the dilation operator (for $N = 1$) is the (self-adjoint) quantization of $p \cdot q$ given by

$$p \cdot q \mapsto \frac{-i\hbar}{2}(\nabla_q q + q\nabla_q).$$

Note that in dimension $d = 3$, the angular momentum $x \wedge p$ vector doesn't require symmetrization and yields the operator

$$x \wedge p \mapsto J := -i\hbar q \wedge \nabla_q, \qquad (24)$$

whose components satisfy the relations

$$[J_i, J_j] = i\hbar J_k, \quad \text{for} \quad (i, j, k) \text{ a permutation of } (1, 2, 3). \qquad (25)$$

The (components of the) angular momentum are unbounded operators which are known to have discrete spectrum, see below.

For a general classical observable $B(q, p)$ (belonging to some reasonable class of smooth real valued functions on $\Gamma \simeq \mathbb{R}^{2d}$, say) the Weyl quantization procedure $B \mapsto B^W$ defined by

$$(B^W \psi)(q) = (2\pi\hbar)^{-d} \int \int B\left(\frac{q + q'}{2}, p\right) e^{i(q-q')\cdot p/\hbar}\psi(q')dq'dp$$

is a good prescription to obtain the corresponding self-adjoint observables. It maps functions of $q \in \mathbb{R}^d$ to the corresponding multiplication operators and polynomials in p_j to the same polynomials in the differential operator $\frac{\hbar}{i}\partial_{q_j}$. Note, however, that there exits other quantization prescriptions that have their own merits.

Also, in other cases, if the geometry of the classical phase space of the system has more structure, the formal operator (22) needs to be supplemented by boundary conditions determined by physical considerations. For example, if the system is confined to a region Λ in configuration space, one customarily provides $\partial\Lambda$ with Dirichlet boundary conditions. In particular, the Hamiltonian of a particle in $\mathbb{R}^d$ confined to a cube Λ whose sides have length L is given by

$$H_\Lambda = -\frac{\hbar^2}{2m}\Delta \quad \text{plus Dirichlet boundary condition at } \partial\Lambda. \qquad (26)$$

P3: *The result of the measure of an observable B on the quantum system characterized by $\psi \in \mathcal{H}$ is an element $b \in \mathbb{R}$ of the spectrum $\sigma(B)$ of the self-adjoint operator B. Moreover, the probability to obtain an element in $(b_1, b_2]$ as the result of this measure on the state ψ is given by*

$$\mathbb{P}_\psi(B \in (b_1, b_2]) = \|P_B((b_1, b_2])\psi\|^2, \qquad (27)$$

where $P_B(I)$ denotes the spectral projector of the operator B on the set $I \subset \mathbb{R}$. Furthermore, once a measure of B is performed, and the result yields a value in a set $(b_1, b_2] \subset \mathbb{R}$, the wave function ψ is reduced, i.e. it undergoes the instantaneous change

$$\psi \mapsto \frac{P_B((b_1, b_2])\psi}{\|P_B((b_1, b_2])\psi\|}. \qquad (28)$$

Another observable C is said to be compatible with B if B and C commute, i.e. if

$$[P_B(\alpha), P_C(\beta)] = 0, \quad \text{for any intervals } \alpha, \beta \subseteq \mathbb{R}.$$

This postulate explains the importance of the efforts made by mathematical physicists in order to determine the spectral properties of operators related to Quantum Mechanics. As, in general, the spectrum of a self-adjoint operator is the union of its discrete and continuous components, the result of the measure of an observable may be quantized, even though its classical counterpart may take values in an interval. This justifies the adjective Quantum for the theory.

Several examples of this fact will be discussed in the lectures on the spectral theory of unbounded operators, in particular for Schrödinger operators of the form $-\Delta + V$. Note that due to the Spectral Theorem, the expectation value of an observable B in a state ψ can be written as

$$\mathbb{E}_\psi(B) = \int_{\sigma(B)} b \|P(db)\psi\|^2 = \int_{\sigma(B)} b\langle\psi|P(db)\psi\rangle = \langle\psi|B\psi\rangle.$$

The reduction processes of the wave function (28) after the measurement of B insures that an immediate subsequent measure of B gives a result in the same set $(b_1, b_2]$ with probability one. The compatibility condition insures that the observables B and C can be simultaneously measured, in the sense that once B and C have been measured with results in the sets α and β respectively, further successive measurements of B and C, in any order, will give results in the same sets with probability one. Or in other words, B and C can be diagonalized simultaneously. This is not the case if B and C do not commute.

Let us introduce some very classical examples as illustrations of the above. The interpretation of the wave function, in the setting where $\Gamma = \mathbb{R}^{2dN}$, is that $|\psi(q)|^2 dq$ is the probability that the system is at point q of configuration space and if $\hat{\psi}(p)$ denotes the Fourier transform of ψ, $|\hat{\psi}(p)|^2 dp$ is the probability that it has momentum p. This is just a particular case of the above rule. Indeed, the operators $q_{k,m}$, $k = 1, \cdots, N$, $m = 1, \cdots, d$ all commute and they have continuous spectra $\mathbb{R}$, as multiplication operators. Hence the interpretation of $|\psi(q)|^2$ follows. That of $|\psi(p)|^2$ is a consequence of the fact that the Fourier transform is unitary on $L^2(\mathbb{R}^{dN})$ which transforms the derivative into a multiplication by the independent variable. Note that (23) shows that $q_{j,n}$ and $p_{j,n}$ cannot be simultaneously determined, an expression of *Heisenberg's uncertainty principle*. Actually, Heisenberg's principle can be put on more quantative grounds as follows. Let A and B be two self-adjoint operators such that their commutator can be written as

$$[A, B] = iC, \quad \text{where } C = C^*.$$

Then, denoting the variance of A in the state ψ by

$$\Delta_\psi(A)^2 = \langle\psi|(A - \mathbb{E}_\psi(A))^2\psi\rangle,$$

we get the inequality

$$\Delta_\psi(A)\Delta_\psi(B) \geq \frac{\mathbb{E}_\psi(C)}{2}. \tag{29}$$

Applied to the operators p and q, we get the familiar relation

$$\Delta_\psi(p)\Delta_\psi(q) \geq \frac{\hbar}{2}.$$

Similarly, the components of the angular momentum J_k, $k = 1, 2, 3$ are not compatible, however it follows from (24) that J_k and $J^2 := J_1^2 + J_2^2 + J_3^2$ are compatible observables, for any k. Hence one can measure the third component of the angular momentum and its length simultaneously. The result of such measures belongs to the spectra of these operators which is discrete as we recall here. If we denote by $\mathcal{K}_j$ the eigenspace of J^2 associated with

the quantum number j and consider the restriction of J_3 to that subspace, a classical algebraic computation shows that

$$\sigma(J^2) = \{j(j+1)|j \in \mathbb{N}/2\} \quad \text{and} \quad \sigma(J_3|_{\mathcal{K}_j}) = \{-j, -j+1, \cdots, j-1, j\}. \quad (30)$$

The effect of boundary conditions on the spectrum of operators can be quite dramatic, as the following comparison shows. The Hamiltonian (22) (with $N = 1$ for simplicity) with $V \equiv 0$, the so-called free Hamiltonian $H_0 = -\frac{\hbar^2}{2m}\Delta$ is unitarily equivalent by Fourier transform to a multiplication operator

$$H_0 \simeq \text{mult } k^2 \quad \text{on } L^2(\mathbb{R}^d).$$

Its spectrum is then $\sigma(H_0) = \mathbb{R}^+$. The spectrum of the operator (26) is easily computed to be

$$\sigma(H_\Lambda) = \{\frac{2\pi^2\hbar^2}{mL^2}(n_1^2 + n_2^2 + \cdots + n_d^2) \,|\, n_j \in \mathbb{Z}\} \quad (31)$$

Another celebrated Hamiltonian is the harmonic oscillator, which will play a prominent role in the quantization of classical fields. In one dimension, this Schrödinger operator reads

$$\frac{p^2}{2m} + \frac{\gamma}{2}q^2, \quad \text{where } \gamma \text{ is a positive constant.}$$

Performing the (canonical) change of operators P and Q by

$$P = (m\gamma)^{-1/4}p, \quad Q = (m\gamma)^{1/4}q, \quad \text{so that} \quad [Q, P] = i\hbar,$$

the operator becomes

$$H_o = \frac{\omega}{2}(P^2 + Q^2), \quad \text{with} \quad \omega = \left(\frac{\gamma}{m}\right)^{1/2}. \quad (32)$$

The spectral analysis of (32) is essentially algebraic once one introduces the *creation* and *annihilation* operators by

$$a^* = \frac{1}{\sqrt{2}}(Q - iP), \quad a = \frac{1}{\sqrt{2}}(Q + iP), \quad \text{such that} \quad [a, a^*] = \mathbb{I}. \quad (33)$$

The operator (32) takes the form

$$H_o = \hbar\omega(a^*a + \frac{1}{2}).$$

Then, defining the vacuum state state $|0\rangle$ as the (normalized) solution to the differential equation

$$a|0\rangle = 0 \quad \Longleftrightarrow \quad |0\rangle = \left(\frac{m\omega}{\hbar\pi}\right)^{1/4} e^{-\frac{m\omega}{2\hbar}q^2},$$

one sets by induction

$$|n\rangle = \frac{(a^*)^n}{\sqrt{n!}}|0\rangle, \;\; n = 1, 2, \cdots$$

These vectors are normalized and take the form of a product of polynomials of degree n, known as Hermite polynomials, by the gaussian $|0\rangle$. One sees easily, using the so-called *canonical commutation relations* (33), that

$$H_o|n\rangle = \hbar\omega(n + \frac{1}{2})|n\rangle \;\; \text{since} \;\; a^*|n\rangle = \sqrt{n+1}|n+1\rangle \;\; \text{and} \;\; a^*|n\rangle = \sqrt{n}|n-1\rangle.$$

These eigenvectors are non-degenerate, such that $\Delta_{|n\rangle}(p)\Delta_{|n\rangle}(q) = \hbar(n + 1/2)$, and span $L^2(\mathbb{R})$.

P4: *The time evolution of the system is determined by its Hamiltonian H, the energy observable of the system. There are two equivalent standard descriptions:*

The Schrödinger picture, in which the state ψ evolves in time according to the time-dependent Schrödinger equation in $\mathcal{H}$

$$i\hbar\frac{d}{dt}\psi(t) = H\psi(t), \quad \text{with} \;\; \psi(0) = \psi. \tag{34}$$

The Heisenberg picture, in which the state is fixed, whereas the observables B evolve in time according to Heisenberg equation in the space of self-adjoint operators on $\mathcal{H}$

$$i\hbar\frac{d}{dt}B(t) = -[H, B(t)], \quad \text{with} \;\; B(0) = B. \tag{35}$$

Introducing the unitary evolution group $U(t) = e^{-itH/\hbar}$ (Spectral Theorem again), we get the relation between the Schrödinger and Heisenberg pictures through the identity

$$\mathbb{E}_{\psi(t)}(B) = \mathbb{E}_\psi(B(t)), \; \forall t \in \mathbb{R},$$

which follows from

$$\psi(t) = U(t)\psi, \;\; \text{and} \;\; B(t) = U(t)^*BU(t).$$

As a consequence, $\psi(t)$ remains normalized for all times and, since the Hamiltonian H commutes with the evolution group $U(t)$ it generates, the observable energy is constant in time. This is also true for observables which are compatible with H, as (35) shows.

The motivations behind the first order linear evolution equation (34) stem from physical observations leading to the so-called superposition principle implying linearity and from the fact that ψ at time 0 should determine completely the state at any later time t. This equation is the quantum equivalent of (3), whereas (35) is the quantum equivalent of (11).

Applied to our example (2), the relation between (35) and (11) together with (9) and (23) are in keeping with the so-called correspondence principle stating that Poisson brackets are to be replaced by commutators in order to achieve formal quantization in that setting:

$$\{\,\cdot\,,\,\cdot\,\} \mapsto \frac{[\,\cdot\,,\,\cdot\,]}{i\hbar}.$$

This yields another motivation for (35).

As an example, consider the case where the Hamiltonian has discrete non-degenerate spectrum $\{E_j\}_{j\in\mathbb{N}}$ with associated eigenvectors $\{\phi_j\}_{j\in\mathbb{N}}$, as is the case if the potential is confining. The time evolution of any initial state ψ reads

$$\psi(t) = \sum_{j\in\mathbb{N}} c_j e^{-itE_j/\hbar}\phi_j, \ \ \forall t\in\mathbb{R}, \ \ \text{where } c_j = \langle\phi_j|\psi\rangle. \tag{36}$$

Therefore, the probability of measuring the energy $E_{j_0} \in \sigma(H)$ in the state $\psi(t)$ is given by

$$\mathbb{P}_{\psi(t)}(H\in\{E_{j_0}\}) = \||\phi_{j_0}\rangle\langle\phi_{j_0}|\psi(t)\|^2$$
$$= |\langle\phi_{j_0}|\psi(t)\rangle|^2 = |c_{j_0}e^{-itE_{j_0}/\hbar}|^2 = |c_{j_0}|^2$$

and is constant. We used the convenient notation $P_H(\{E_j\}) = |\phi_j\rangle\langle\phi_j|$. Similarly, the probability to obtain an energy in a subset $\mathcal{E} = \{E_{j_0}, E_{j_n}, \cdots, E_{j_n}\}$ of the spectrum of H is given by

$$\mathbb{P}_{\psi(t)}(H\in\mathcal{E}) = \left\|\sum_k |\phi_{j_k}\rangle\langle\phi_{j_k}|\psi(t)\right\|^2 = \sum_k^n |c_{j_k}|^2. \tag{37}$$

Note however, that the sole knowledge of the spectrum of H does not allow in general to get precise information about the evolution of states that are not eigenstates, due to the complicated interferences present in (36). A nice and sometimes useful exception to this rule is the case of *coherent states* for the harmonic oscillator. In the one-dimensional setting used in (32), these normalized states depend on a complex number α and are defined as

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}}|n\rangle.$$

They have the properties (which can be checked by means of (33) only)

$$a|\alpha\rangle = \alpha|\alpha\rangle, \ \ \Delta_{|\alpha\rangle}(p)\Delta_{|\alpha\rangle}(q) = \hbar/2, \ \ |\alpha\rangle|_{\alpha=0} = |0\rangle.$$

Their explicit expression as functions of $L^2(\mathbb{R})$ reads as

$$|\alpha\rangle = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left(-\frac{1}{2}\alpha(\alpha+\alpha^*) - \frac{m\omega}{2\hbar}q^2 + \left(\frac{2m\omega}{\hbar}\right)^{1/2}\alpha q\right).$$

Then, using $e^{-itH/\hbar}|n\rangle = e^{-i\omega(n+1/2)t}|n\rangle$, we get

$$e^{-itH/\hbar}|\alpha\rangle = e^{-i\omega t/2}|e^{i\omega t}\alpha\rangle.$$

Finally, we note also here that in case the Hamiltonian is time-dependent, (34) gives rise under some regularity conditions to a two-parameter unitary evolution operator $U(t,s)$ on $\mathcal{H}$ satisfying

$$\frac{\partial}{\partial t}U(t,s) = H(t)U(t,s), \;\; \text{with} \;\; U(s,s) = I\!I$$

from which follows the relation

$$U(t,r)U(r,s) = U(t,s), \;\; \forall r,s,t \in \mathbb{R}.$$

In such a case, the evolution operator is no longer an exponential and the future of initial wave functions is usually harder to describe. But again, in case the Hamiltonian is essentially a quadratic form in p and q, with time-dependent coefficients, explicit solutions to the above equation can be obtained, provided the initial conditions are of a coherent state type.

1.3 Fermions and Bosons

So far we have considered systems which have no internal structure or degrees of freedom. Such internal degrees of freedom are introduced by taking a tensor product of the original Hilbert space $\mathcal{H}$ with $\mathcal{K}$, another Hilbert space in which these degrees of freedom live, so that the system is now described by means of the new Hilbert space $\mathcal{H} \otimes \mathcal{K}$. An important internal degree of freedom is the spin of a particle. It is a vector valued operator S in a finite dimensional space $\mathcal{K} \simeq \mathbb{C}^k$ whose components also satisfy the commutation relations (25), and therefore displays the same spectral properties (30). A spin is half-integer or integer, depending whether it is true for the maximal quantum number s of S^2.

Consider now the state of a collection of N identical particles, that is sharing the same physical characteristics like masses, spins, charge, etc. In the framework of our example and slightly abusing notations, it is described by means of a wave function $\psi(q_1, s_1; q_2, s_2; \cdots; q_N, s_N) \in L^2(\mathbb{R}^{dN}) \otimes \mathcal{K}^N$. The fact that the particles are identical is equivalent to saying that all observables $B(q_1, p_1, s_1; q_2, p_2, s_2; \cdots, q_N, p_N, s_N)$ applied to such states are invariant under permutations of their variables (q_j, p_j, s_j). An example of such observable is the kinetic energy part of (22), and it is also true of the potential part of this Hamiltonian due to (2). Therefore, if P_{jk} is the operator whose action is to permute the variables labeled j and k, one has

$$P_{jk}^2 = 1, \;\; P_{jk}^* = P_{jk}, \;\; \text{and} \;\; [P_{jk}, B] = 0, \;\; \forall \; B. \tag{38}$$

Hence, $\sigma(P_{jk}) = \{+1, -1\}$ and the observables and P_{jk} can be diagonalized simultaneously. Thus we can first diagonalize P_{jk} and then describe the

observables restricted to the corresponding subspaces of P_{jk}. It is a law of nature that the eigenvalues of P_{jk} are either +1 for all pairs j, k or -1 for all pairs j, k. Therefore, identical particles divide themselves into two distinct sets of particles: those that are invariant under exchange of the variables of their wave function, and those that undergo a sign change under this operations. Particles belonging to the former set are called *bosons* whereas they are called *fermions* if they belong to the latter. As the properties (38) are true for all pairs jk of labels, they are also true for arbitrary permutations $\pi \in \mathcal{P}_N$ in the group of permutations of N elements. In particular, if P_π denotes the permutation of indices corresponding to π, then $P_\pi \psi = \psi$ for bosons and $P_\pi \psi = (-1)^\pi \psi$ for fermions, where $(-1)^\pi$ is the signature of the permutation π. In other words, the above discussion shows that the physical Hilbert spaces for bosons and fermions are not the N-fold tensor product $\mathcal{H}^N$ of the Hilbert space $\mathcal{H}$ of their one particle descriptions, but rather $\mathcal{H}^n_+$ and $\mathcal{H}^n_-$, defined by

$$\mathcal{H}^n_\pm = \mathcal{S}_\pm \mathcal{H}^n \ \text{where} \ S_\pm = \frac{1}{N!} \sum_{\pi \in \mathcal{P}_N} (\pm 1)^\pi P_\pi.$$

The operators $S_\pm$ are easily checked to be orthogonal projectors onto the orthogonal subspace $\mathcal{H}^n_\pm$ of $\mathcal{H}^n$ and vectors in these spaces are characterized by the properties described above. These characteristics will have important consequences on the physical properties of collections of such particles. In particular, antisymmetry forbids two independent fermions to be in the same quantum state. Indeed, the sign change induced by exchange of these two particles in the antisymmetrization procedure makes the vector vanish. This goes under the name *Pauli's Principle*.

It turns out that the fermionic or bosonic nature of particles is linked to the properties of their spin. Indeed, it can be shown within the realm of relativistic quantum field theory that fundamental requirements of Physics as micro-causality and Lorentz's invariance imply the so-called *Spin-Statistic Theorem* asserting that particles with half-integer spin are fermions, whereas those with integer spin are bosons. It is an experimental fact that no other statistics is present in nature. Electrons are thus fermions and photons are bosons (although the latter have an internal degree of freedom called helicity, instead of a spin).

2 Quantum Statistical Mechanics

2.1 Density Matrices

We give here, in a particular setting, some heuristics behind the formal definition of *state* (or mixed state) in Quantum Statistical Mechanics which will be used later on. The first approach is based on a time dependent point of view

supplemented by a postulate on the behaviour of some phases, advocated in [H], for example.

Let us start from (37) which says that in case the Hamiltonian H on $\mathcal{H}$, of a system is time independent, discrete and non-degenerate, *i.e.* when the system is isolated from the rest of the world, its normalized wave function can be written as

$$\psi(t) = \sum_{j \in \mathbb{N}} c_j(t) \phi_j, \tag{39}$$

with explicit time dependent complex valued coefficients $c_j(t)$. The basic postulates of Statistical Mechanics are formulated for isolated systems. However, in practice, one is often interested in a subsystem of the whole system only, so that, certain degrees of freedom are not observed and are incorporated in what we call the rest of the world. Thus Statistical Mechanics effectively deals with systems that interact with the external world, so that the truly isolated system is our initial system plus the rest of the world. The relevant Hilbert space to describe this new system is the tensor product $\mathcal{R} \otimes \mathcal{H}$, where $\mathcal{R}$ is the Hilbert space of the rest of the world, and the corresponding scalar product is the product of the respective scalar products on $\mathcal{R}$ and $\mathcal{H}$. The wave function of this larger isolated system can still be written as (39), with the proviso that the $c_j's$ are now time dependent elements of $\mathcal{R}$ such that $\sum_j \langle c_j(t) | c_j(t) \rangle_{\mathcal{R}} = 1$, where the subscript specifies what scalar product is used.

Now, if B is an observable acting on the original system only, technically of the form $I\!\!I \otimes B$ as an operator on $\mathcal{R} \otimes \mathcal{H}$, the instantaneous expectation value of a set of measurements of this observable on $\psi(t)$ is given by

$$\langle \psi(t) | I\!\!I \otimes B \psi(t) \rangle_{\mathcal{R} \otimes \mathcal{H}} = \sum_{k,j} \langle c_j(t) | c_k(t) \rangle_{\mathcal{R}} \langle \phi_j | B \phi_k \rangle_{\mathcal{H}}.$$

In an actual experiment, it is rather the time average of the above quantity that is measured, over a time that is large with respect to the "molecular time scale", but short with respect to the resolution of the measurement apparatus. Therefore, the measured quantity is actually

$$\langle B \rangle = \sum_{k,j} \overline{\langle c_j(t) | c_k(t) \rangle}_{\mathcal{R}} \langle \phi_j | B \phi_k \rangle_{\mathcal{H}},$$

where the bar indicates time average.

One postulate concerns the scalar products of the $c_j(t)$'s about which we have minimal knowledge, since it deals with properties of the external world. It is postulated that these scalar products producing interferences are averaged to zero over the time scale on which we observe the system, this is the *Random phases* postulate:

$$\overline{\langle c_j(t) | c_k(t) \rangle}_{\mathcal{R}} = 0, \ \ \forall \ j \neq k.$$

The consequence of this postulate is an effective description of the original system by means of the non-zero scalars

$$\lambda_j = \overline{\langle c_j(t)|c_j(t)\rangle}_{\mathcal{R}}, \quad \text{such that} \quad \sum_j \lambda_j = 1, \tag{40}$$

in the sense that the outcome of a measurement in that framework is given by

$$\langle B \rangle = \sum_j \lambda_j \langle \phi_j | B\phi_j \rangle_{\mathcal{H}}.$$

Therefore one says that the random phases postulate allows to regard the state of the system as an incoherent superposition of eigenstates of H or a *mixed state*.

As a consequence, it is possible to represent the mixed state by a *density matrix*. Let ρ be the linear operator on $\mathcal{H}$ defined by

$$\rho = \sum_j \lambda_j |\phi_j\rangle\langle\phi_j|. \tag{41}$$

It is a positive, trace class operator, of trace one, such that

$$\langle B \rangle = \text{Tr } (\rho B).$$

This operator contains all the information we have about the mixed state and allows to compute in a convenient way all expectation values by means of the trace operation.

Note that a pure state χ as defined in the previous section corresponds to the density matrix $\rho_\chi = |\chi\rangle\langle\chi|$. Actually, it is easy to see that a density matrix ρ corresponds to a pure state if and only if it is a rank one projector.

Another approach of mixed states consists in noting that incomplete information about a system always leads to density matrices, without resorting to delicate properties of the time evolution.

A first point of view consists in considering the system as an *ensemble* of true eigenstates, to be considered one at a time, where the relative proportion of the eigenstate ϕ_j is λ_j. The value $\lambda_j \in [0,1]$ is interpreted as the classical probability to get the pure eigenstate ϕ_j in the mixed state. This statistical interpretation $\rho = \sum_j \lambda_j |\phi_j\rangle\langle\phi_j|$ allows to avoid any consideration of effective coupling between the system and the "external world" and makes no use of *a priori* knowledge about the time evolution.

A second interpretation of incomplete knowledge about the system consists in splitting the total Hilbert space it lives in into $\mathcal{R} \otimes \mathcal{H}$, where $\mathcal{R}$ concerns the degrees of freedom that are not known. Therefore, if $\{\varphi_j \otimes \psi_k\}_{j,k}$ denotes an orthonormal basis of $\mathcal{R} \otimes \mathcal{H}$ made out of individual orthonormal bases of $\mathcal{R}$ and $\mathcal{H}$ and ρ is any density matrix on $\mathcal{R} \otimes \mathcal{H}$, one introduces the corresponding reduced density matrix $\rho_{\mathcal{H}}$ on $\mathcal{H}$ by its matrix elements

$$(\rho_{\mathcal{H}})_{ij} = \sum_k \langle \varphi_k \otimes \psi_i \,|\rho\; \varphi_k \otimes \psi_j \rangle, \quad \forall i, j.$$

The matrix $\rho_{\mathcal{H}}$ is designed so that for any operator of the form $I\!\!I \otimes B$, one has

$$\text{Tr}_{\mathcal{H}}(\rho_{\mathcal{H}} B) = \text{Tr}\,(\rho\; I\!\!I \otimes B),$$

where $\text{Tr}_{\mathcal{H}}$ denotes the trace in $\mathcal{H}$. This formula is in keeping with our ignorance of the degrees of freedom in $\mathcal{R}$ which are traced out. The point is that one can easily check that if $\rho = |\Psi\rangle\langle\Psi|$ for some $\Psi \in \mathcal{R} \otimes \mathcal{H}$, then, in general, the corresponding reduced density matrix $\rho_{\mathcal{H}}$ characterizes a mixed state.

Therefore we will adopt from now on the following definition:

A mixed state (or simply state) in Quantum Statistical Mechanics is a positive trace class operator on $\mathcal{H}$ of trace 1.

The time dependence of density matrices is governed by the following equation in the subset of density matrices in $\mathcal{T}(\mathcal{H})$, the linear space of trace class operators on $\mathcal{H}$,

$$i\hbar \frac{d}{dt}\rho = [H, \rho], \;\; \rho(0) = \rho_0. \tag{42}$$

This equation stems from the cyclicity of the trace and the relation which must hold for all t and all observables B

$$\text{Tr}\,(\rho B(t)) = \text{Tr}(\rho(t)B).$$

Note that (42) is in keeping with the evolution of the density matrix of a pure state.

3 Boltzmann Gibbs

So far, we haven't talked about equilibrium properties. The basic postulate in Quantum Statistical Mechanics describes the density matrix of an isolated system that has reached equilibrium.

Assume the energy of the system is known to lie within the range $[E, E + \Delta]$, where $\Delta << E$. The *Equal a priori Probability* postulate, formalizing again our minimal knowledge of the total system, states within the framework where the ϕ_j's represent eigenstates of the Hamiltonian H, that at equilibrium, the λ_j's defining the density matrix ρ_{eq} are given by

$$\lambda_j = \begin{cases} \lambda & \text{if } E < E_j < E + \Delta \\ 0 & \text{otherwise} \end{cases} \tag{43}$$

The constant λ is normalized so that the trace of ρ_{eq} is one. Other ways of writing ρ_{eq} are

$$\rho_{eq} = P_H(E, E + \Delta)/\operatorname{Tr}(P_H(E, E + \Delta)) = \frac{\sum_{E < E_j < E < \Delta} |\phi_j\rangle\langle\phi_j|}{\#\{j \,|\, E_j \in]E, E + \Delta[\}}. \quad (44)$$

Note that as ρ_{eq} is a function of the Hamiltonian H, it is constant in time due to (42), which is what we expect for the equilibrium density matrix.

This above prescription for ρ_{eq} corresponds to the *micro-canonical* ensemble, where it is understood that the system under consideration has fixed energy and fixed number of particles.

In this case, Boltzmann's formula is used to define the *entropy*

The entropy of a system at equilibrium in the microcanonical ensemble reads

$$S(E) = k \ln \Gamma(E), \quad where \ \ \Gamma(E) = \#\{j \,|\, E_j \in]E, E + \Delta[\} \quad (45)$$

and $k \simeq 1,38 \times 10^{-23} J/K$ is Boltzmann's constant.

The quantity $\Gamma(E)$, which is the denominator in (44), gives the number of quantum states that are accessible to the system. The entropy actually depends on other variables such as the volume V of the system, the number N of particles of the system, etc... that we omitted in the notation. The definition (45) makes the bridge between equilibrium Statistical Mechanics and thermodynamics, once the thermodynamic limit is taken. That is, once it is demonstrated that Boltzmann's formula fulfills the following conditions:
a) extensivity, so that the thermodynamical limit as $V \to \infty, N \to \infty, E \to \infty$ exists i.e.

$$\frac{1}{N}S(E, V, N) \to s(e, v), \quad where \ \ E/N \to e, \ V/N \to v, \ S/N \to s$$

with e, s and v are the densities of energy and entropy and v is the specific volume.
b) the fact that if exterior parameters of the system initially at equilibrium are varied in such a way that the system can reach another equilibrium configuration, then the difference of entropies between these configurations is non negative. This is an expression of the second law of thermodynamics which implies that the equilibrium state maximizes the entropy. We'll come back to this point shortly.

Without discussing thermodynamics, we mention that assuming the existence of entropy and that properties a) and b) hold, all thermodynamical quantities can be computed from $S(E, V, N)$ through the definitions:

$$\frac{1}{T} = \left.\frac{\partial S}{\partial E}\right)_{V,N} \quad \text{defines the } temperature, \text{ so that },$$

$$U(S, V, N) \equiv E(S, V, N) \quad \text{the } internal \ energy \text{ of the system exists and}$$

$$P = -\frac{\partial U}{\partial V}\bigg)_{S,N} \quad \text{defines the } \textit{pressure} \text{ whereas}$$

$$\mu = \frac{\partial U}{\partial N}\bigg)_{S,V} \quad \text{defines the } \textit{chemical potential.}$$

The above definitions yield the familiar differential

$$dU(S,V,N) = TdS - PdV + \mu dN, \tag{46}$$

motivating the physical interpretations of these derivatives. The extensivity property of U, i.e. homogeneity of degree one, associated with (46) implies

$$U = TS - PV + \mu N. \tag{47}$$

In order to complete the picture, let us briefly recall that the first law of thermodynamics asserts that the differential

$$dU = \delta Q - \delta W \quad \text{is exact,}$$

where δQ is amount of heat absorbed by the system and δW is the work done by the system in any transformation. A corollary of the second law of thermodynamics says that the differential

$$dS = \frac{\delta Q}{T} \quad \text{is exact,}$$

relating the experimental notion of heat to entropy. These two statements imply the existence of the functions S and U at equilibrium.

It can be argued that the definition (45) satisfies requirement a), but we shall not provide the argument here. Let us consider b). This last property calls for a variational approach of the entropy. Hence we introduce a more general definition of the entropy of a state by

The entropy of state ρ of a physical system is given by

$$S(\rho) = -k \, Tr\left(\rho \ln(\rho)\right), \tag{48}$$

where the function $x \mapsto x\ln(x)$ is defined to be zero at $x = 0$.

If $\{\lambda_j\}_{j \in N}$ denotes the set of eigenvalues of the density matrix

$$\rho = \sum_j \lambda_j |\phi_j\rangle\langle\phi_j|,$$

the entropy is given by

$$S(\rho) = -k \sum_j \lambda_j \ln(\lambda_j). \tag{49}$$

Therefore one sees that $S(\rho) \geq 0$ with $S(\rho) = 0$ if and only if $\lambda_j = \delta_{j,k}$ for some k, i.e. ρ corresponds to a pure state. Also, the entropy (48) of the density matrix ρ describing two independent systems defined on $\mathcal{H}_1$ and $\mathcal{H}_2$ is the sum of the individual entropies, as expected. Indeed, in such a case, $\rho = \rho_1 \otimes \rho_2$ on $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, where ρ_j, $j = 1, 2$ are the density matrices of the individual systems. Using

$$\ln(\rho_1 \otimes \rho_2) = \ln(\rho_1) \otimes I\!I + I\!I \otimes \ln(\rho_2),$$

and taking partial traces, one gets $S(\rho) = S(\rho_1) + S(\rho_2)$. More generally, it can be shown that for $\alpha \in [0, 1]$ and arbitrary density matrices ρ_1, ρ_2

$$S(\alpha \rho_1 + (1 - \alpha)\rho_2) \geq \alpha S(\rho_1) + (1 - \alpha)S(\rho_2),$$

which shows that S is concave as a function on the set of density matrices, i.e. mixing density matrices increases the entropy. And, on the other hand, the entropy is almost convex in the sense

$$S(\alpha \rho_1 + (1 - \alpha)\rho_2) \leq \alpha S(\rho_1) + (1 - \alpha)S(\rho_2) - \alpha \ln \alpha - (1 - \alpha)\ln(1 - \alpha).$$

More mathematical properties of the entropy are provided in [BR] or [S], for example.

Now, maximizing (49) over the probabilities λ_j's shows that the entropy is maximal when the eigenvalues are all constant. Thus, when the corresponding eigenvectors ϕ_j are those of the Hamiltonian, we get back both the equal a priori probability postulate and Boltzmann's formula.

The micro canonical ensemble is convenient to motivate definitions, but one often prefers to use the *canonical* ensemble for applications. In that setting, the system under consideration interacts with a thermal reservoir whose property is to remain at fixed temperature. Exchanges of energy are allowed between the reservoir and the system, under the constraint that the averaged energy of the system is kept fixed. The maximization of entropy in the canonical ensemble leads to Gibbs prescription for the density matrix, as we now (formally) argue.

Consider the functional over the set of density matrices.

$$\mathcal{F}(\rho) = S(\rho) - k\beta \langle H \rangle_\rho, \tag{50}$$

where $\langle H \rangle_\rho$ denotes the expectation value of the energy computed by means of the density matrix ρ, and β is a Lagrange multiplier associated with the energy constraint.

We need to compute the first variation of $\mathcal{F}$

$$\delta \mathcal{F}(\rho) = \frac{d}{dt}\mathcal{F}(\rho + t\eta)|_{t=0},$$

where the admissible variation η is any trace class operator of zero trace and $\operatorname{Tr}(\rho) = 1$. In order to do so, we first justify the following intuitive relation:

If $A(t)$ is a t-dependent self-adjoint operator, such that $A(t) = A(0) + t\eta$, then

$$\frac{d}{dt}\operatorname{Tr}(f(A(t))) = \operatorname{Tr}(f'(A(t))\eta). \tag{51}$$

Indeed, Hellman-Feynman formula applied to $A(t)$ whose eigenvalues and normalized eigenvectors are denoted by $(a_j(t), \varphi_j(t))$ reads

$$a_j'(t) = \langle \varphi_j(t) | A'(t)\varphi_j(t) \rangle. \tag{52}$$

Thus for any (reasonable) real valued function f,

$$\frac{d}{dt}\operatorname{Tr}(f(A(t))) = \sum_j f'(a_j(t))a_j'(t) = \sum_j \langle \varphi_j(t) | f'(A(t))\varphi_j(t) \rangle a_j'(t). \tag{53}$$

In the case under consideration, $A'(t) = \eta$ so that, by (52), $a_j'(t) = \langle \varphi_j(t) | \eta\varphi_j(t) \rangle$ and, using orthonormality of the φ_j's, the RHS of (53) equals

$$\sum_j \langle \varphi_j(t) | f'(A(t))\varphi_j(t) \rangle \langle \varphi_j(t) | \eta\varphi_j(t) \rangle$$

$$= \sum_{j,k} \langle \varphi_j(t) | f'(A(t))\varphi_k(t) \rangle \langle \varphi_k(t) | \eta\varphi_j(t) \rangle$$

$$= \sum_j \langle \varphi_j(t) | f'(A(t))\eta\varphi_j(t) \rangle, \quad \text{which yields (51).}$$

In our case, we get

$$\delta\mathcal{F}(\rho) = -k\operatorname{Tr}(\eta(\ln(\rho) + I\!I + \beta H)),$$

which has to be zero for any admissible η if ρ extremalizes $\mathcal{F}$. In particular, we can choose $\eta = \sum_j \eta_j |\varphi_j\rangle\langle\varphi_j|$ where $\{\varphi_j\}$ are the set of eigenvectors of the self adjoint operator $\ln(\rho) + I\!I + \beta H$ and $\{\eta_j\}$ are a set of real numbers satisfying $\sum_j \eta_j = 0$. For *that* η, we have

$$\delta\mathcal{F}(\rho) = -k\sum_j \eta_j v_j,$$

where the v_j's denote the eigenvalues of $\ln(\rho) + I\!I + \beta H$. Choosing above $\eta_0 = -\eta_1 \neq 0$ and $\eta_j = 0$ if $j \geq 2$, we get that $\delta\mathcal{F}(\rho) = 0$ for any η_0 implies $v_0 = v_1$. Iterating, we get $v_0 = v_j$, for any $j \in \mathbb{N}$. Hence

$$\delta\mathcal{F}(\rho) = 0 \ \forall\eta \iff \ln(\rho) + I\!I + \beta H = v_0 I\!I.$$

The constant v_0 will be determined by the normalization of the state. Therefore, exponentiating, we find as extremalizer the *Gibbs distribution*

$$\rho_G = \frac{e^{-\beta H}}{\mathrm{Tr}\left(e^{-\beta H}\right)}. \tag{54}$$

Explicit computations that we do not present here show that the second variation $\delta^2 \mathcal{F}$ is negative. Hence we get that the Gibbs prescription yields a maximum of the functional (50). Here the parameter β used to insure constancy of the average energy of the system in that canonical ensemble setting will be identified with the inverse temperature given by the usual formula

$$\beta = \frac{1}{kT}. \tag{55}$$

The normalization of the Gibbs distribution

$$Z := \mathrm{Tr}\left(e^{-\beta H}\right)$$

defines the (canonical) *partition function*. It is related to the internal energy of the system $\langle H \rangle_{\rho_G}$ by

$$\langle H \rangle_{\rho_G} = -\frac{\partial \ln(Z)}{\partial \beta}.$$

Again, in the thermodynamic limit, the partition function of Statistical Mechanics is directly linked to a thermodynamical quantity: the free energy F of the system, defined as $F = U - TS$ (see below), where U is the internal energy $\langle H \rangle_{\rho_G}$ of the system. To substanciate this claim, let us formally compute by means of (54) (assuming the thermodynamic limit and extensivity holds),

$$S = k\beta \langle H \rangle_{\rho_G} + k \ln(Z) \quad \text{so that}$$
$$F := -kT \ln Z \tag{56}$$

defines the *free energy* in Statistical Mechanics.

Moreover, as a consequence of our variational approach, we get that the free energy of a system is minimized by the equilibrium state, which together with (56) are two familiar properties of the free energy.

More precisely, in thermodynamics, F is a function of (T, V, N), which is the result of its very definition:

The thermodynamical free energy is the Legendre transform of the internal energy $U(S, V, N)$ with respect to the variable S

$$F(T, V, N) = (U - TS)(T, V, N) \tag{57}$$

where $S(T, V, N)$ is computed from $\left.\frac{\partial U}{\partial S}\right)_{V,N} = T$.

This operation allows to trade the entropy variable for the more natural temperature variable. Recall that the Legendre transform of a one variable function $f : x \mapsto f(x)$ is the function $a : p \mapsto a(p)$ defined by

$$a(p) = f(x(p)) - px(p),$$

where $x(p)$ is obtained by inversion of

$$\frac{d}{dx}f(x) = p(x).$$

There is no loss of information in the process as long as the inversion of $f'(x)$ is possible and when f is concave, respectively convex, its Legendre transform is convex, respectively concave. In case f depends on other variables y, one has the identities

$$\frac{\partial}{\partial p}a(p, y) = -x(p, y), \quad \frac{\partial}{\partial y}a(p, y) = \frac{\partial}{\partial y}f(x, y)$$

allowing to recover all thermodynamic quantities from F via the relations

$$\left.\frac{\partial F}{\partial T}\right)_{V,N} = -S, \quad \left.\frac{\partial F}{\partial V}\right)_{T,N} = -P, \quad \left.\frac{\partial F}{\partial N}\right)_{T,V} = \mu.$$

We can now provide a justification of the identification (55) as follows, assuming the thermodynamic limit is taken and extensivity holds. Indeed, by means of that identification, we get by explicit computation on (56)

$$\frac{\partial}{\partial T}F = -k\ln Z - Tk\frac{\text{Tr}\left(\frac{\partial}{\partial\beta}e^{-\beta H}\right)}{Z}\frac{\partial\beta}{\partial T} = -k\ln Z - \frac{\langle H\rangle_{\rho_G}}{T} = -S,$$

which is identical to the first relation above.

Let us present here the classical computation of partition functions associated with independent harmonic oscillators.

Let $\mathcal{H}$ be the Hilbert space spanned by the eigenvectors $\{|n\rangle\}_{n=0,\cdots,\infty}$ of the Hamiltonian H_o in (32) corresponding to the energies $\epsilon_n = \frac{1}{2} + n$, $n = 0, \cdots, \infty$ (we assume $\hbar\omega = 1$, without loss). Working in the canonical ensemble, the partition function of one harmonic oscillator reads

$$Z_1(\beta) = \text{Tr}\left(e^{-\beta H_o}\right) = e^{-\beta/2}\sum_{n=0}^{\infty}e^{-\beta n} = \frac{1}{e^{\beta/2} - e^{-\beta/2}},$$

the internal energy $U_1(\beta) = \langle H_o\rangle_{\rho_G}$ is given by

$$U(\beta) = -\frac{\partial\ln(Z_1(\beta))}{\partial\beta} = \frac{1}{2}\coth(\beta/2),$$

whereas the free energy $F_1(\beta)$ reads

$$F(\beta) = -kT\ln(Z_1(\beta)) = kT\ln(e^{\beta/2} - e^{-\beta/2}).$$

In case we work with a d-dimensional harmonic oscillator, or, equivalently, with d independent oscillators with the same frequency, we denote by $|n_1, n_2, \cdots, n_d\rangle$, $n_j \in \mathbb{N}$, $j = 1, \cdots, d$, the eigenvector corresponding to the energy $\frac{d}{2} + n_1 + \cdots + n_d$. Thus, the corresponding partition function reads

$$Z(\beta, d) = e^{-\beta d/2} \sum_{n_1 \geq 0, n_2 \geq 0, \cdots, n_d \geq 0} e^{-\beta(n_1 + \cdots + n_d)} = Z(\beta)^d,$$

so that the internal and free energies and are given by

$$U(\beta, d) = dU(\beta), \quad F_d(\beta, d) = dF(\beta). \tag{58}$$

Going from the micro canonical to the canonical ensemble, we have allowed energy exchanges between the system under consideration and a thermal reservoir. In a similar fashion, we can relax the condition that the number of particles in the system is fixed and allow particles exchanges with the reservoir as well, assuming the their average number only is fixed. This corresponds to working in the *grand canonical ensemble*. As we will see later on, allowing particles exchanges in Quantum Open Systems is essential, in the sense that the statistical properties of these particles, i.e. their bosonic or fermionic nature, have definite physical consequences.

This calls for a precision about the Hilbert space suitable to describe such situations, the so-called *second quantization formalism*. The Hilbert space allowing variable numbers of particles is either the symmetrical or anti-symmetrical *Fock space*, depending on the statistics. These Hilbert spaces will be object of much mathematical care later on, so we will briefly and informally describe here the bosonic and fermionic Fock spaces $\mathcal{F}_\pm(\mathcal{H})$. If $\mathcal{H}$ is the one-particle Hilbert space, the n-fold properly symmetrized tensor product $\mathcal{H}_\pm^n$ is the n-boson or n-fermion subspace. An element Ψ of $\mathcal{F}_\pm(\mathcal{H})$ is a collection $\{\psi(n)\}_{n \in \mathbb{N}}$, where $\psi(n) \in \mathcal{H}_\pm^n$, for all $n > 0$, $\psi(0) \in \mathbb{C} \equiv \mathcal{H}_\pm^0$, with the obvious linear structure and norm $\|\Psi\|_\pm^2 = \sum_n \|\psi(n)^2\|_\pm$. Observables $\mathbf{B}$ on the Fock space can be constructed as $\mathbf{B} = \sum_n B(n)$, where the $B(n)$'s acting on the n-particle subspaces are given, (with $B(0) = 0$). In particular, the number operator $\mathbf{N}$ defined by $\mathbf{N}\Psi = \{n\psi(n)\}_{n \in \mathbb{N}}$ has the form $\mathbf{N} = \sum_n n \, I\!I_{\mathcal{H}_\pm}$. Another case is that of *one body operators*. That is when $\mathbf{A} = \sum_n A(n)$ with $A(n) = \sum_{j=1}^n A_j$, where $A_j = I\!I \otimes \cdots I\!I \otimes A \otimes I\!I \cdots I\!I$, and A acts on the j'th copy of $\mathcal{H}$.

With these preliminaries behind us, let us assume we are given a Hamiltonian $\mathbf{H}$ with the structure above. The equilibrium state in that framework is obtained by maximization of the entropy, under the constraints that both the average energy and average number of particles are fixed. This leads to the computation of the first and second variations of the functional $\mathcal{G}$ over the density matrices defined by

$$\mathcal{G}(\rho) = S(\rho) - k\beta\langle \mathbf{H}\rangle_\rho + k\beta\mu\langle \mathbf{N}\rangle_\rho,$$

where β and μ are Lagrange multipliers associated with the imposed constraints. They will be identified in the thermodynamic limit, with the inverse temperature and chemical potential, respectively. A maximizing procedure quite similar to the one performed above that we will not detail here yields the extremum

$$\rho_{GC} = \frac{e^{-\beta(\mathbf{H} - \mu\mathbf{N})}}{\mathcal{Z}}$$

where, due to the structure of $\mathbf{H}$ and $\mathbf{N}$, the *grand canonical partition function* $\mathcal{Z}$ can be written as

$$\mathcal{Z} = \sum_n e^{\beta\mu n} \operatorname{Tr}_{\mathcal{H}_\pm^n}(e^{-\beta H(n)}).$$

The quantity $z = e^{\beta\mu}$ is also called the *fugacity* and with Z_n the canonical partition function, we can rewrite

$$\mathcal{Z} = \sum_n z^n Z_n.$$

One can also verify that the maximal value of $\mathcal{G}$ is

$$S(\rho_{GC}) - k\beta\langle\mathbf{H}\rangle_{\rho_{GC}} + k\beta\mu\langle\mathbf{N}\rangle_{\rho_{GC}} = k\ln\mathcal{Z}. \tag{59}$$

To make the bridge with thermodynamics, consider the thermodynamical *grand potential* Φ defined by the Legendre transform of F with respect to N, i.e.

$$\Phi(T, V, \mu) = (F - \mu N)(T, V, \mu)$$

where $\frac{\partial F}{\partial N} = \mu$. One can then see by formal manipulations similar to those performed above, assuming the thermodynamic limit and extensivity holds, that Φ is minimal at equilibrium. From (59) and (47), we get that this minimum is given by

$$\Phi = -PV = -kT\ln\mathcal{Z}$$

and we further have the thermodynamical relations

$$\left.\frac{\partial\Phi}{\partial T}\right)_{V,\mu} = -S, \quad \left.\frac{\partial\Phi}{\partial\mu}\right)_{T,V} = -N.$$

The ensemble (microcanonical, canonical or grand canonical) chosen to describe a specific system is largely made according to convenience for the computations. Therefore it is comforting to know that the respective descriptions are all equivalent. This is the statement known as the *equivalence of ensembles* which says that in the thermodynamical limit, one can use either the microcanonical or the canonical ensemble to perform calculations of thermodynamical quantities because the results will agree. Instead of providing a justification of this statement here, we shall be content with the explicit

verification of this fact for a system of independent harmonic oscillators considered in the microcanonical and canonical ensembles.

In the microcanonical ensemble, we compute the entropy by means of (45). If we have N independent oscillators each of which has energy levels $j + 1/2$, $j \in \mathbb{N}$, we get for N large,

$$\Gamma(E) \simeq \# \left\{ n_j \in \mathbb{N}, j = 1, \cdots, N, \ | \ \sum_j n_j = E - N/2 \right\} \simeq \binom{E + N/2}{N},$$

using the combinatoric formula

$$\# \left\{ n_j \in \mathbb{N}, j = 1, \cdots, N, \ | \ \sum_j n_j = M \in \mathbb{N} \right\} = \binom{M + N - 1}{N - 1}.$$

Hence, by means of Stirling formula, we compute in term of the energy density $e = E/N$,

$$S(E, N) \simeq Nk \left((e + \frac{1}{2}) \ln(e + \frac{1}{2}) - (e - \frac{1}{2}) \ln(e - \frac{1}{2}) \right) \equiv Ns(e).$$

Therefore, the temperature is determined by

$$\frac{1}{T} = \frac{\partial s(e)}{\partial e} = k \ln \left(\frac{e + \frac{1}{2}}{e - \frac{1}{2}} \right),$$

so that we get the following formula for the energy density

$$e = \frac{1}{2} \left(\frac{e^\beta + 1}{e^\beta - 1} \right) = \frac{1}{2} \coth(\beta/2).$$

For the same system considered in the canonical ensemble, we obtained in (58) with $d = N$,

$$U(\beta, N) = NU(\beta) = N \frac{1}{2} \coth(\beta/2),$$

which yields the same energy density $e = U/N$.

Let us consider now the computation of the grand canonical partition function, in the simple bosonic/fermionic context where particles do not interact with one another.

Consider the normalized vector $|n_0, n_1, n_2, \cdots, n_j, \cdots\rangle_\pm \in \mathcal{F}_\pm(\mathcal{H})$ in the so-called *occupation number representation* relative to the eigenstates states $|n\rangle$ in $\mathcal{H}$ of some nondegenerate hamiltonian H. This vector consists in a normalized, fully (anti)symmetrized tensor product of states $|n\rangle \in \mathcal{H}$ characterized by n_0 factors $|0\rangle$, n_1 factors $|1\rangle$, $\cdots n_j$ factors $|j\rangle$, etc. The number of particles N in such a state is obviously given by $N = \sum_k n_k$. In

case of bosons, $n_j \in \mathbb{N}$ without restriction, whereas in case of fermions, Pauli's principle enforces $n_j \in \{0, 1\}$, for any j. We'll denote by $\mathbb{N}'$ the set of allowed values of the n_j's , depending on the statistics. The collection $\{|n_0, n_1, n_2, \cdots, n_j, \cdots\rangle_\pm\}_{n_0 \in \mathbb{N}', \cdots, n_j \in \mathbb{N}'}$ forms an orthonormal basis of $\mathcal{F}_\pm(\mathcal{H})$. If one considers only the states with a fixed number of particles, one gets that the set $\{|n_0, n_1, n_2, \cdots, n_j, \cdots\rangle_\pm \mid \sum_k n_k = N \}$ forms an orthonormal basis of the subspace $\mathcal{H}_\pm^N$.

Let ϵ_n denote the eigenvalue of H corresponding to $|n\rangle$. Then the one body observable $\mathbf{H}$ in $\mathcal{F}_\pm(\mathcal{H})$ constructed from H satisfies

$$\mathbf{H}\, |n_0, n_1, n_2, \cdots, n_j, \cdots\rangle_\pm = \sum_k n_k \epsilon_k \, |n_0, n_1, n_2, \cdots, n_j, \cdots\rangle_\pm.$$

The corresponding physical system consists of a collection of independent fermions or bosons individually driven by the Hamiltonian H. Though quite simple, such systems allow to put forward the effect of the statistics. The canonical partition function $Z_N(\beta)$ of N independent fermions/bosons is

$$Z_N(\beta) = \sum_{\{n_j| \sum_j n_j = N\}} e^{-\beta \sum_j n_j \epsilon_j},$$

where the restrictions on the n_j's due to the statistics are implicit in the notation. Hence, with $z = e^{\beta\mu}$,

$$\mathcal{Z}(\beta, z) = \sum_{N \geq 0} z^N \sum_{\{n_j| \sum_j n_j = N\}} e^{-\beta \sum_j n_j \epsilon_j} = \sum_{N \geq 0} \sum_{\{n_j| \sum_j n_j = N\}} \prod_j (z e^{-\beta\epsilon_j})^{n_j}$$

$$= \prod_j \left[\sum_n (z e^{-\beta\epsilon_j})^n \right] = \begin{cases} \prod_j (1 - z e^{-\beta\epsilon_j})^{-1} & \text{for bosons} \\ \prod_j (1 + z e^{-\beta\epsilon_j}) & \text{for fermions.} \end{cases}$$

In particular, we compute

$$\langle N \rangle_{\rho_{GC}} = z \frac{\partial}{\partial z} \ln(\mathcal{Z}(\beta, z)) = \begin{cases} \sum_j \frac{z e^{-\beta\epsilon_j}}{1 - z e^{-\beta\epsilon_j}} & \text{for bosons} \\ \sum_j \frac{z e^{-\beta\epsilon_j}}{1 + z e^{-\beta\epsilon_j}} & \text{for fermions,} \end{cases}$$

which allows to determine z. Similarly, the average occupation numbers can be obtained as

$$\langle n_j \rangle_{\rho_{GC}} = -\frac{1}{\beta} \frac{\partial}{\partial \epsilon_j} \ln(\mathcal{Z}(\beta, z)) = \frac{z e^{-\beta\epsilon_j}}{1 \mp z e^{-\beta\epsilon_j}} \quad \begin{cases} \text{for bosons} \\ \text{for fermions.} \end{cases}$$

Therefore, we get the expected relation

$$\langle N \rangle_{\rho_{GC}} = \sum_j \langle n_j \rangle_{\rho_{GC}},$$

where we clearly see the effects of the statistics and temperature on the the average occupation numbers.

References

[BR] O. Bratteli, D. W. Robinson, *Operator Algebras and Quantum Statistical Mechanics II*, Texts and Monographs in Physics, Springer, New York, Heidelberg, Berlin, 1981

[GJ] J. Glimm, A. Jaffe, *Quantum Physics*, Springer, New York, Heidelberg, Berlin, 1981

[H]   K. Huang, *Statistical Mechanics*, J. Wiley & Sons, New York, London, Sydney, 1963

[MR] Ph. A. Martin, F. Rothen, *Many -body Problems and Quantum Field Theory*, Texts and Monographs in Physics, Springer, 2nd Edition, 2004

[S]   B. Simon, *The Statistical Mechanics of Lattice Gases*, Princeton Series in Physics, Princeton New-Jersey, 1993