Chapter 2

The standard formulations of classical and quantum mechanics

I first start by reminders of classical mechanics, probabilities and quantum mechanics, in their usual formulations in theoretical physics. This is mostly very standard material. The last section on reversibility and probabilities in quantum mechanics is a slightly more original presentation of these questions.

2.1 Classical mechanics

Classical mechanics can be formulated using the Lagrangian formulation or the Hamiltonian formulation. They apply both to non relativistic systems of particles, to fields (like the electromagnetic field) and to relativistic systems. They are valid for closed non-dissipative systems. Macroscopic systems with dissipation, and more generally out of equilibrium open systems must be studied by the tools of statistical mechanics. This is still very active field of research, but out of the scope of this short presentation.

The standard books on classical mechanics are the books by Landau & Lifshitz [LL76] and the book by A. Arnold [AVW89].

2.1.1 The Lagrangian formulation

In the Lagrangian formulation, a classical (at this point non-relativistic) system is described by its configuration space $\mathcal{C}$ (a Lagrangian manifold). A point $\mathbf{q} = \{q^i\}_{i=1,N}$ of $\mathcal{C}$ describes an instantaneous configuration of the system. The N coordinates q^i label the N degrees of freedom (d.o.f.) of the system. The state of the system at time t is given by its configuration $\mathbf{q}$ and its velocity $\dot{\mathbf{q}} = \frac{d\mathbf{q}}{dt}$ in configuration space. The dynamics is given by the equations of motion, that takes the form of the Euler-Lagrange equations. They derive from the Lagrange function (the Lagrangian) $L(\mathbf{q}, \dot{\mathbf{q}})$ and read

$$\frac{d}{dt} \frac{\partial L(q, \dot{q})}{\partial \dot{q}(t)} = \frac{\partial L(q, \dot{q})}{\partial q(t)} \tag{2.1.1}$$

Let us consider the simple textbook case of a non relativistic particle of mass m in a one dimensional space (a line). Its coordinate (position) is denoted q. It is submitted to a conservative force which derives from a potential $V(q)$. The potential is independent of time. The velocity is $\dot{q}(t) = \frac{dq}{dt}$. The dynamics of the particle is given by Newton's equation

$$m\,\ddot{q}(t) = -\frac{\partial}{\partial q}\,V(q) \tag{2.1.2}$$

which is obtained from the Lagrangian

$$L(q, \dot{q}) = \frac{m}{2}\dot{q}^2 - V(q) \tag{2.1.3}$$

The equation of motion derives from the least action principle. The action $S[q]$ of a general trajectory $q = \{q(t)\}$ starting from q_1 at time t_1 and ending at q_2 at time t_2 is defined as the integral of the Lagrangian

$$S[q] = \int_{t_i}^{t_f} dt\, L(q(t), \dot{q}(t)) \tag{2.1.4}$$

The classical trajectory q_c, that satisfy the equations of motion, is the trajectory that extremizes the action S under small variations that leaves the initial and final points unmoved.

$$q(t) = q_c(t) + \delta q(t)\,, \quad \delta q(t_1) = \delta q(t_2) = 0 \tag{2.1.5}$$

The stationarity condition

$$q(t) = q_c(t) + \delta q(t)\,, \quad \delta q(t_i) = \delta q(t_f) = 0 \quad \Longrightarrow \quad S[q_c + \delta q] = S[q_c] + \mathcal{O}(\delta q^2) \tag{2.1.6}$$

can be rewritten as the vanishing of the functional derivative of the action

$$\frac{\delta S[q]}{\delta q(t)} = -\frac{d}{dt}\frac{\partial L(q, \dot{q})}{\partial \dot{q}(t)} + \frac{\partial L(q, \dot{q})}{\partial q(t)} = 0 \tag{2.1.7}$$

which leads to the Euler-Lagrange equations and to Newton's equation 2.1.2.

The Lagrangian formalism is valid for most conservative physical systems. For N particles in d dimensions the configuration space is $N \times d$ dimensional, and a configuration is a $d \times N$-dimensional vector $\mathbf{q} = \{q_a^i; i = 1, d, a = 1, N\}$. The formalism is more generally valid for charged particles in an external magnetic field, for mechanical systems with constraints and submitted to external forces, for the classical electromagnetic field and for general classical fields, provided that the dynamics is reversible and non-dissipative (but the dynamics may be non-invariant under time reversal, think about the well known example of the motion of a charged particle in a magnetic field). It applies also to relativistic systems and relativistic fields, and to the space-time of general relativity as well (with a proper treatment of age symmetries, of space and time and of Lorentz and diffeomorphism invariance)

The Lagrangian formulation is especially well suited to discuss the role and consequences of symmetries. For instance the relativistic scalar field (the classical Klein-Gordon field) is described by a classical real field $x \to \phi(x)$ in Minkowski space-time
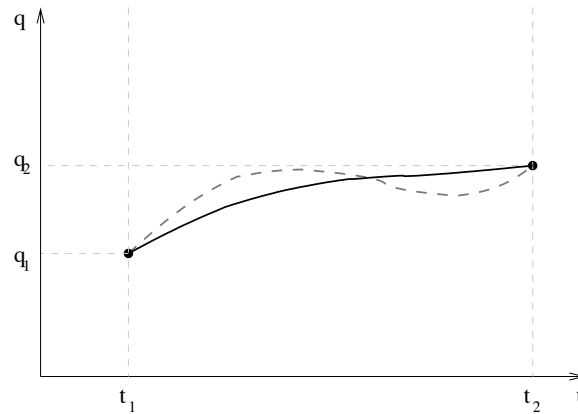
Figure 2.1: The least action principle: the classical trajectoire (full line) extremizes the action. Under a variation (dashed line) $\delta S = 0$. The initial and final positions are kept fixed.

$x = (x^\mu) = (t, \vec{x}) \in \mathbb{M}^{1,3}$, with metric $h = \mathrm{diag}(-1, 1, 1, 1)$, i.e. space-time line element $ds^2 = -dt^2 + d\vec{x}^2$. Its Lagrangian density L and its action S are

$$L(\phi, \partial\phi) = -\frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{m^2}{2}\phi^2 \quad , \qquad S[\phi] = \int d^4x\, L(\phi, \partial\phi) \qquad (2.1.8)$$

and are explicitly invariant under Lorentz and Poincaré transformations.

It is in this Lagrangian formalism that it is easy to prove the famous Noether's theorem. This theorem states that to any symmetry (continuous group of invariance) of the dynamics (the action) is associated a conserved charge (a conserved local current for a field).

The "teleological aspect" of the least action principle (the trajectory is defined as a function of its initial and final conditions) seemed for a long time mysterious and had to wait for quantum mechanics and its path integral formulation by R. Feynman to be fully understood. The Lagrangian formalism is thus well suited to discuss path integral and functional integral quantization.

2.1.2 The Hamiltonian formulation

2.1.2.a - The phase space and the Hamiltonian

The Hamiltonian formulation is in general equivalent to the Lagrangian formulation, but in fact slightly more general. It is well suited to discuss the relation between classical physics and the "canonical quantization" schemes. It allows to discuss in general the structure of the states and of the observables of a classical system, without reference to a specific choice of configuration variables and of configuration space.

For a classical system with n degrees of freedom, the set of possible states of the system is its phase space Ω. A state of the system is now a point x in the phase space Ω of the system. Ω is a manifold with even dimension $2n$. The evolution equations

are flow equations (first order differential equations in time) in the phase space, and derive from the Hamiltonian function $H(\mathbf{x})$.

For the particle in dimension $d = 1$ in a potential there is one degree of freedom, $n = 1$, so that $\Omega = \mathbb{R}^2$ and $\dim(\Omega)=2$. The two coordinates in phase space are the position q et the momentum p of the particle.

$$\mathbf{x} = (q, p) \tag{2.1.9}$$

The Hamiltonian is

$$H(q, p) = \frac{p^2}{2m} + V(q) \tag{2.1.10}$$

The equations of motion are the Hamilton equations

$$\dot{p} = -\frac{\partial H}{\partial q} \quad , \quad \dot{q} = \frac{\partial H}{\partial p} \tag{2.1.11}$$

The relation between the momentum and the velocity $p = m\dot{q}$ is now a dynamical relation. The dissymmetry between q and p comes in fact from the antisymmetric Poisson bracket structure present in the phase space, this will be discussed a bit later.

The Hamilton equations derive also from a variational principle. To find the classical trajectory in phase space such that $q(t_1) = q_1$, $q(t_2) = q_2$ one has to extremize the action functional (the "Hamiltonian action") $\mathcal{S}_H$

$$\mathcal{S}_H[q, p] = \int_{t_1}^{t_2} dt \left(p(t)\dot{q}(t) - H(q(t), p(t)) \right) \tag{2.1.12}$$

with respect to variations of $q(t)$ and of $p(t)$, $q(t)$ being fixed at the initial and final times $t = t_1$ et t_2, but $p(t)$ being left free at $t = t_1$ and t_2. Indeed, the functional derivatives of $\mathcal{S}_H$ are

$$\frac{\delta \mathcal{S}_H}{\delta q(t)} = -\dot{p}(t) - \frac{\partial H}{\partial q}(q(t), p(t)) \,, \quad \frac{\delta \mathcal{S}_H}{\delta p(t)} = \dot{q}(t) - \frac{\partial H}{\partial p}(q(t), p(t)) \tag{2.1.13}$$

The change of variables $(q, \dot{q}) \to (q, p)$, of the Lagrangian to the Hamiltonian $L(q, \dot{q}) \to H(q, p)$ and of the action functionals $S[q, \dot{q}] \to \mathcal{S}_H[q, p]$ between the Lagrangian and the Hamiltonian formalism correspond to a Legendre transformation with respect to the velocity $\dot{q}$. The velocity $\dot{q}$ and the momentum p are conjugate variables. Indeed one has the relation

$$\mathbf{p} = \frac{\partial L(\mathbf{q}, \dot{\mathbf{q}})}{\partial \dot{\mathbf{q}}} \quad , \quad H(\mathbf{q}, \mathbf{p}) = \mathbf{p}\dot{\mathbf{q}} - L(\mathbf{q}, \dot{\mathbf{q}}) \tag{2.1.14}$$

2.1.2.b - Hamilton-Jacobi equation

For a classical trajectory $q_{\rm cl}(t)$ solution of the equations of motion, the "Hamiltonian" action $\mathcal{S}_H[q_c, p_c]$ and the Lagrangian action $S[q_c]$ are equal. This is not really surprising, this is a property of the Legendre transformation.

Now let us fix the initial time t_1 and the initial position q_1 of the particle. This classical action can be considered now as being a function of the final time $t_2 = t$ and
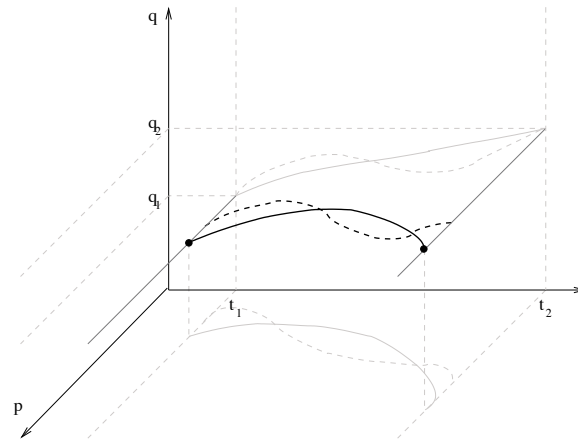
Figure 2.2: The least action principle in phase space: The classical trajectory (full line) extremizes the action $S_H[q,p]$. The initial and final positions are fixed. The initial and final momenta are free. Their actual value is given by the variational principle, and is a function of the initial and final positions and of the initial and final times.

of the final position $q(t_2) = q_2 = q$ of the particle. This function is called the Hamilton-Jacobi action, or the Hamilton function (not to be confused with the Hamiltonian), and let us denote it $\mathcal{S}(q,t) = \mathcal{S}_{\mathrm{HJ}}(q,t)$ to be explicit (the initial conditions $q(t_1) = q_1$ being implicit)

$$\mathcal{S}(q,t) = \mathcal{S}_{\mathrm{HJ}}(q,t) = \mathcal{S}[q_{\mathrm{cl}}] \quad \text{with } q_{\mathrm{cl}} \text{ classical solution such that} \quad q(t_2) = q, t_2 = t$$
$$\text{and where } t_1 \text{ and } q(t_1) = q_1 \text{ are kept fixed} \qquad (2.1.15)$$

Using the equations of motion it is easy to see that the evolution with the final time t of this function $\mathcal{S}(q,t)$ is given by the differential equation

$$\frac{\partial \mathcal{S}}{\partial t} = -H\left(q, \frac{\partial \mathcal{S}}{\partial q}\right) \qquad (2.1.16)$$

with H the Hamiltonian function. This is is a first order differential equation with respect to the final time t. It is called the Hamilton-Jacobi equation.

From this equation on can show that (the initial conditions (t_1, q_1) being fixed) the impulsion p and the total energy E of the particle, expressed as a function of its final position q and of the final time t, are

$$E(q,t) = -\frac{\partial \mathcal{S}}{\partial t}(q,t), \quad p(q,t) = \frac{\partial \mathcal{S}}{\partial q}(q,t) \qquad (2.1.17)$$

These equations extend to the case of systems with n degrees of freedom and of more general Hamiltonians, positions and momenta being now n components vectors

$$\mathbf{q} = \{q^i\} \quad , \qquad \mathbf{p} = \{p_i\} \qquad i = 1, \cdots, n \qquad (2.1.18)$$

The Hamilton-Jacobi equations are quite important, in particular in the context of the semi-classical limit of quantum mechanics, where the Hamilton-Jacobi action turns to be the "quantum phase" of the quantum particle.

2.1.2.c - Symplectic manifolds

For the most general case (for example of classical systems with constraints), the Hamiltonian formulation requires the formalism of symplectic geometry. The phase space Ω of a system with n degrees of freedom is a manifold with an even dimension $N = 2n$. It is not necessarily the Euclidean space $\mathbb{R}^N$. As we shall see it is only locally like $\mathbb{R}^N$, but it may have a non trivial topology leading to interesting physical effects.

Locally the phase space Ω is described by local coordinates $\mathbf{x} = \{x^i, i = 1, 2n\}$. Warning! now the $2n$ coordinates x_i are coordinate in phase space, not some physical spatial position coordinates in configuration space. Again in general such coordinates systems are not global and must be patched together by some coordinate changes (with properties to be defined later) to get a global description of the phase space.

The Hamiltonian dynamics requires a symplectic structure on Ω. This symplectic structure allows to define (or amounts to define) the Poisson brackets for the system. Ω is said to be a symplectic manifold if it is embodied with an antisymmetric 2-form ω (a degree 2 differential form) which is non-degenerate and closed ($d\omega = 0$). This means that to each point $\mathbf{x} \in \Omega$ is associated (in the coordinate system $\{x^i\}$ the $2n \times 2n$ antisymmetric matrix $\mathbb{R}^N$, that defines the 2-form

$$\omega(\mathbf{x}) = \frac{1}{2}\omega_{ij}(\mathbf{x})\, dx^i \wedge dx^j \tag{2.1.19}$$

We use here the standard notations of differential geometry. $dx^i \wedge dx^j$ is the antisymmetric product (exterior product) of the two 1-forms dx^i and dx^j.

The matrix $\omega(\mathbf{x})$ is antisymmetric and non-degenerate,

$$w_{ij}(\mathbf{x}) = -w_{ji}(\mathbf{x}) \quad , \qquad \det(\omega) \neq 0 \tag{2.1.20}$$

This implies that it is invertible. The form ω is moreover closed. This means that its exterior derivative $d\omega$ is zero

$$d\omega(\mathbf{x}) = 0 \tag{2.1.21}$$

with

$$d\omega(\mathbf{x}) = \frac{1}{3!}\sum_{i,j,k} \partial_i \omega_{jk}(\mathbf{x})\, dx^i \wedge dx^j \wedge dx^k \tag{2.1.22}$$

In term of components this means

$$\forall\, i_1 < i_2 < i_3 \quad , \qquad \partial_{i_1}\omega_{i_2 i_3} + \partial_{i_2}\omega_{i_3 i_1} + \partial_{i_3}\omega_{i_1 i_2} = 0$$

The fact that ω is a differential form means that under a local change of coordinates $\mathbf{x} \to \mathbf{x}'$ (in phase space) the components of the form change as

$$\mathbf{x} \to \mathbf{x}' \quad , \qquad \omega = \omega(\mathbf{x})_{ij}\, dx^i \wedge dx^j = \omega'(\mathbf{x}')_{ij}\, dx'^i \wedge dx'^j$$

that is

$$\omega'(\mathbf{x}')_{ij} = \omega(\mathbf{x})_{kl}\, \frac{\partial x^k}{\partial x'^i}\frac{\partial x^l}{\partial x'^j}$$

The symplectic form allows an intrinsic definition of the Poisson brackets (see below).

2.1.2.d - Simple example

For the simple case of a particule on a ligne, $n = 1$, $\Omega = \mathbb{R}^2$, $\mathbf{x} = (q, p)$, The symplectic form is simply $\omega = dq \wedge dp$. Its components are

$$\omega = (\omega_{ij}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{2.1.23}$$

For a particle in n dimensional space, the configuration space is $\mathbb{R}^n$, the phase space is $\Omega = \mathbb{R}^{2n}$, with coordinates $\mathbf{x} = (q^i, p^i)$, and the symplectic form is $\omega = \frac{1}{2} \sum_i dq^i \wedge dp^i$. It can be written as the block matrix

$$(\omega_{ij}) = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ -1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ 0 & 0 & -1 & 0 & \cdots \\ & & & & \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{2.1.24}$$

Darboux's theorem states that for any symplectic manifold (Ω, ω), it is always possible to find locally a coordinate systems such that the symplectic form takes the form 2.1.24. In such a Darboux coordinate system ω is constant and is a direct sum of antisymmetric symbols. The (q^i, p^j) are said to be local pairs of conjugate variables (the generalization of the conjugate position and momentum variables).

The fact that locally the symplectic form may be written under this generic constant form means that symplectic geometry is characterized by global invariants, not by local ones. This is different from Riemaniann geometry, where the metric tensor g_{ij} cannot in general be written as a flat metric $h_{ij} = \delta_{ij}$, and where there are local invariants, such as the scalar curvature R, and many others.

2.1.2.e - Observables, Poisson brackets

The observables of the system defined by a symplectic phase space (Ω, ω) are identified with the "smooth" real functions from $\Omega \to \mathbb{R}$ (smooth means "sufficiently regular", at least differentiable and in general C^∞) real functions from $\Omega \to \mathbb{R}$. The (observed) value of an observable f for the system in the state $\mathbf{x}$ is simply the value of the function $f(\mathbf{x})$. Of course observables may depend also explicitly on the time t.

$$\text{system in state } \mathbf{x} \quad \to \quad \text{value of } f = f(\mathbf{x}) \tag{2.1.25}$$

For two differentiable functions (observables) f and g, their Poisson bracket $\{f, g\}_\omega$ is the function defined by

$$\{f, g\}_\omega(x) = \omega^{ij}(x)\, \partial_i f(x) \partial_j g(x) \quad \text{with} \quad \partial_i = \frac{\partial}{\partial x^i} \quad \text{and} \quad w^{ij}(x) = \left(w^{-1}(x)\right)_{ij} \tag{2.1.26}$$

the matrix elements of the inverse of the antisymmetric matrix $\omega(x)$ (remember that ω is non degenerate, hence invertible). When no ambiguity is present, the subscript ω will be omitted in the Poisson bracket $\{f, g\}$.

In a canonical local coordinate system (Darboux coordinates) the Poisson bracket takes the standard form

$$\{f,g\} = \sum_i \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p^i} - \frac{\partial f}{\partial p^i}\frac{\partial g}{\partial q^i} \qquad \text{and} \qquad \{q^i, p^j\} = \delta_{ij} \qquad (2.1.27)$$

The Poisson bracket is bi-linear, antisymmetric

$$\{f,g\} = -\{g,f\} \qquad (2.1.28)$$

Since it involves only first order derivatives iit satisfies the Leibnitz rule (the Poisson bracket acts as a derivation w.r.t to one of its term onto the other one)

$$\{f,gh\} = \{f,g\}h + g\{f,h\} \qquad (2.1.29)$$

The fact that the symplectic form is closed $d\omega = 0$ is equivalent to the Jacobi identity

$$\{f,\{g,h\}\} + \{g,\{h,f\}\} + \{h,\{f,g\}\} = 0 \qquad (2.1.30)$$

Finally, in general coordinates, the knowledge of the Poisson bracket $\{\,,\}$ is equivalent to the knowledge of the symplectic form ω via its inverse ω^{-1}

$$\{x^i, x^j\} = \omega^{ij}(\mathbf{x}) \qquad (2.1.31)$$

2.1.2.f - The dynamics and Hamiltonian flows

The purpose of this formalism is of course to describe the dynamics in phase es-pace. It is generated by an Hamiltonian function H. The Hamiltonian is a real "smooth" (at least differentiable) function on phase space $\Omega \to \mathbb{R}$, like the observables. The state of the system $\mathbf{x}(t)$ evolves with time according to the Hamilton equation

$$\dot{\mathbf{x}}(t) = \{\mathbf{x}(t), H\} \qquad (2.1.32)$$

In terms of coordinates this reads

$$\dot{x}^i(t) = w^{ij}(\mathbf{x}(t))\partial_j H(\mathbf{x}(t)) \qquad (2.1.33)$$

The Hamilton equations involve the Poisson Bracket and is covariant under local changes of coordinates in phase space. They are flow equations of the form

$$\dot{x}^i(t) = F^i(\mathbf{x}(t)) \qquad (2.1.34)$$

but the vector field $F^i = \omega^{ij}\partial_j H$ is very special and derives from the Hamiltonian H. The time flow, i.e. the application $\phi\colon \Omega \times \mathbb{R} \to \Omega$ is called the Hamiltonian flow associated to H. The evolution functions $\phi_t(\mathbf{x})$ defined by

$$\mathbf{x}(t = 0) = \mathbf{x} \quad \Longrightarrow \quad \mathbf{x}(t) = \phi_t(\mathbf{x}) \qquad (2.1.35)$$

form a group of transformations (as long as H is independent of the time)

$$\phi_{t_1 + t_2} = \phi_{t_1} \circ \phi_{t_2} \qquad (2.1.36)$$

More generally, let us consider a (time independent) observable f (a function on Ω). The evolution of the value of the observable f for a dynamical state $\mathbf{x}(t)$, $f(\mathbf{x}, t) = f(\mathbf{x}(t))$ where $\mathbf{x}(t) = \phi_t(\mathbf{x})$, obeys the equation

$$\frac{\partial f(\mathbf{x}, t)}{\partial t} = \{f, H\}(\mathbf{x}(t)) \qquad (2.1.37)$$

where the r.h.s. is the Poisson bracket of the observable f and the Hamiltonian H. In particular (when H is independent of t) the energy $E(t) = H(\mathbf{x}(t))$ is conserved

$$\frac{\partial E(\mathbf{x}, t)}{\partial t} = 0 \qquad (2.1.38)$$

2.1.2.g - The Liouville measure

The symplectic form ω defines an invariant volume element $d\mu$ on phase space Ω.

$$d\mu(\mathbf{x}) = \omega^n = \prod_{i=1}^{2n} dx^i |\omega|^{1/2} \quad , \qquad |\omega| = |\det(\omega_{ij})| \qquad (2.1.39)$$

This is the Liouville measure on Ω. This mesure is invariant under all the Hamitonian flows, and is in fact the only local invariant.

2.1.2.h - A less trivial example: the classical spin

The simplest example of a system with a non trivial phase space is the classical spin (the classical top with constant total angular momentum). The states of the spin are labelled by unit 3-components vector $\vec{n} = (n_1, n_2, n_3)$, $|\vec{n}| = 1$ (the direction of the angular momentum). Thus the phase space is the 2-dimensional unit sphere and is now compact (hence different from $\mathbb{R}^2$)

$$\Omega = \mathcal{S}_2$$

The classical precession equation

$$\frac{d\vec{n}}{dt} = \vec{B} \times \vec{n}$$

can be written in Hamiltonian form. $\vec{B}$ is a vector in $\mathbb{R}^3$, possibly a 3-component vector field on the sphere depending on $\vec{n}$.

There is a symplectic structure on Ω. It is related to the natural complex structure on $\mathcal{S}_2$ (the Riemann sphere). The Poisson bracket of two functions f and g on $\mathcal{S}_2$ is defined as

$$\{f, g\} = (\vec{\nabla} f \times \vec{\nabla} g) \cdot \vec{n}.$$

The gradient field $\vec{\nabla} f$ of a function f on the sphere is a vector field tangent to the sphere, so $\vec{\nabla} f \times \vec{\nabla} g$ is normal to the sphere, hence collinear with $\vec{n}$. In spherical coordinates

$$\vec{n} = (\sin\theta \cos\phi, \sin\theta \sin\phi, \cos\theta)$$

the Poisson bracket is simply

$$\{f, g\} = \frac{1}{\sin\theta}\left(\frac{\partial f}{\partial\theta}\frac{\partial g}{\partial\phi} - \frac{\partial g}{\partial\theta}\frac{\partial f}{\partial\phi}\right)$$

Admissible local Darboux coordinates $\mathbf{x} = (x^1, x^2)$ such that $\omega = dx^1 \wedge dx^2$ must be locally orthogonal, area preserving mappings $\mathbb{R} \to S_2$. Examples are the "action-angle" variables (the Lambert cylindrical equal-area projection)

$$\mathbf{x} = (\cos\theta, \phi)$$

and the plane coordinates (the Lambert azimuthal equal-area projection).

$$\mathbf{x} = (2\sin(\theta/2)\cos\phi, 2\sin(\theta/2)\sin\phi)$$

The Hamiltonian which generates the precession dynamics is simply (for constant $\vec{B}$)
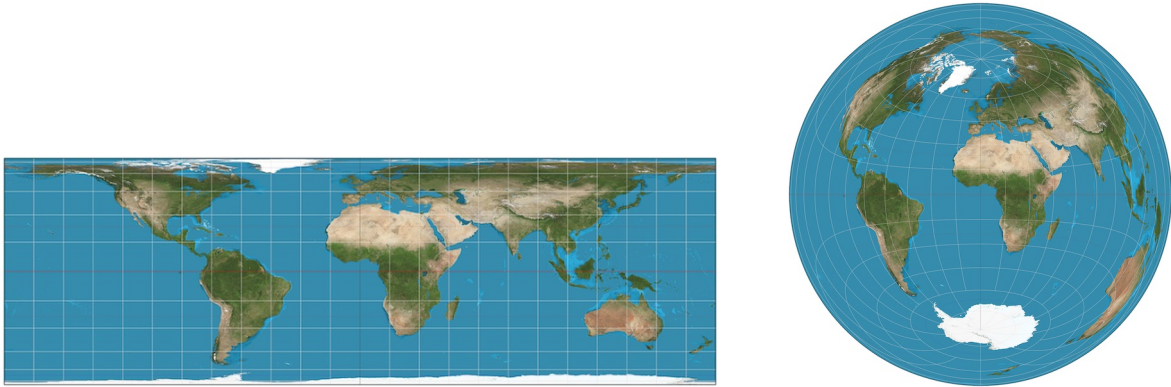

Figure 2.3: The Lambert cylindrical and azimuthal coordinates

$$H = \vec{B} \cdot \vec{n}$$

2.1.2.i - Statistical states, distribution functions, the Liouville equation

We now consider statistical ensembles. If one has only partial information on the state of the system (for instance if one consider an instance of the system taken at random in a big ensemble, or if we know only the value of some of its observables) this information is described by a statistical state (also called mixed) state or statistical ensemble) φ. Mixed states are described by probability distributions on the phase space Ω,

$$d\rho_\varphi(\mathbf{x}) = d\mu(\mathbf{x})\rho_\varphi(\mathbf{x}) \tag{2.1.40}$$

with $d\mu(\mathbf{x})$ the Liouville measure and $\rho_\varphi(\mathbf{x})$ the probability density, a non negative distribution (function) such that

$$\rho_\varphi(\mathbf{x}) \geq 0 \quad , \quad \int_\Omega d\mu(\mathbf{x})\rho_\varphi(\mathbf{x}) = 1 \tag{2.1.41}$$

For a system in a statistical state φ the expectation for an observable f (its expectation value, i.e. its mean value if we perform a large number of measurements of f on independent systems in the same ensemble φ) is

$$\langle f \rangle_\varphi = \int_\Omega d\mu(\mathbf{x}) \, \rho_\varphi(\mathbf{x}) \, f(\mathbf{x}) \tag{2.1.42}$$

When the system evolves according to the Hamiltonian flow ϕ_t generated by an Hamiltonian H, the statistical state evolves with time $\varphi \to \varphi(t)$, as well as the distribution function $\rho_\varphi \to \rho_{\varphi(t)}$. φ being the initial state of the system at time $t = 0$, we denote this distribution function

$$\rho_{\varphi(t)}(\mathbf{x}) = \rho_\varphi(\mathbf{x}, t) \tag{2.1.43}$$

It is related to the initial distribution function at time $t = 0$ by

$$\rho_\varphi(\mathbf{x}(t), t) = \rho_\varphi(\mathbf{x}) \quad , \qquad \mathbf{x}(t) = \phi_t(\mathbf{x}) \tag{2.1.44}$$

(the Liouville measure is conserved by the Hamiltonian flow). Using the evolution equation for $\mathbf{x}(t)$ 2.1.33, one obtains the Liouville equation as the evolution equation for the distribution function $\rho_\varphi(\mathbf{x}, t)$

$$\frac{\partial}{\partial t} \rho_\varphi = \left\{ H, \rho_\varphi \right\} \tag{2.1.45}$$

The entropy of the statistical state φ is given by the Boltzmann-Gibbs formula (with $k_B = 1$)

$$S = - \int_\Omega d\mu(\mathbf{x}) \, \rho_\varphi(\mathbf{x}) \, \log(\rho_\varphi(\mathbf{x})) \tag{2.1.46}$$

Of course when the state of the system is fully determined, it is a "pure state" $\mathbf{x}_0$ ($\varphi_{\text{pure}} = \mathbf{x}_0$) the distribution function is a Dirac measure $\rho_{\text{pure}}(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}_0)$.

2.1.2.j - Canonical transformations

Hamiltonian flows are canonical transformations, i.e. examples of (bijective) mappings $\mathcal{C} \ \Omega \to \Omega$ that preserve the symplectic structure. Denoting $\mathbf{X} = \mathcal{C}(\mathbf{x})$ the image of the point $\mathbf{x} \in \Omega$ by the canonical transformation $\mathcal{C}$, this means simply that the symplectic form ω^* defined by

$$\omega^*(x) = \omega(X) \tag{2.1.47}$$

is equal to the original form $\omega^* = \omega$. ω^* is called the pullback of the symplectic form ω by the mapping $\mathcal{C}$. In a local coordinate system such that $\mathbf{x} = (x^i)$ and $\mathbf{X} = (X^k)$, this means in components

Canonical transformations preserve the Poisson brackets. Let f and g be two observables (functions $\Omega \to \mathbb{R}$ and $F = f \circ \mathcal{C}^{-1}$ and $G = g \circ \mathcal{C}^{-1}$ their transform by the transformation $\mathcal{C}$

$$f(\mathbf{x}) = F(\mathbf{X}) \quad , \qquad g(\mathbf{x}) = G(\mathbf{X}) \tag{2.1.48}$$

$\mathcal{C}$ is a canonical transformation if

$$\{f, g\}_\omega = \{F, G\}_\omega \tag{2.1.49}$$

Taking for f and g the coordinate change $x^i \to X^i$ itself, canonical transformations are change of coordinates such that

$$\{X^i, X^j\} = \{x^i, x^j\} \qquad (2.1.50)$$

Canonical transformations are very useful tools in classical mechanics. They are the classical analog of unitary transformations in quantum mechanics.

In the simple example of the classical spin, the canonical transformations are the smooth area preserving diffeomorphisms of the 2 dimensional sphere.

2.1.2.k - Along the Hamiltonian flows

As an application, one can treat the Hamiltonian flow ϕ_t as a time dependent canonical transformation (a change of reference frame) and consider the dynamics of the system in this new frame, that evolves with the system. In this new coordinates, that we denote $\bar{x} = \{\bar{x}^i\}$, if at time $t = 0$ the system is in the initial state $\bar{x} = x_0$, at time t it is still in the same state $\bar{x}(t) = x_0$. while the observables f become time dependent.

Indeed, starting from a time independent observable f, $x \to f(x)$, let us denote $\bar{f}$ the time dependent observable in the new frame,

$$\bar{f}(\bar{x}, t) = f(x(t)) \qquad \text{with} \quad x(t) = \phi_t(\bar{x}) \qquad (2.1.51)$$

It describes how the observable f evolves with t, as a function of the initial state $\bar{x}$ at $t = 0$. Its evolution is given by

$$\frac{\partial \bar{f}}{\partial t} = \{\bar{f}, H\} \qquad (2.1.52)$$

This change of frame corresponds to changing from a representation of the dynamics by an evolution of the states, the observables being time independent, to a representation where the states do not evolves, but where there observables depend on time. This is the analog for Hamiltonian dynamics to what is done in fluid dynamics: going from the Eulerian specification (the fluid moves in a fixed coordinate system) to the Lagrangian specification (the coordinate system moves along the fluid). These two representations are of course the classical analog of the Schrödinger picture (vector states evolves, operators are fixed) and of the Heisenberg picture (vector states are fixed, operators depend on time) in quantum mechanics.

2.1.3   The algebra of classical observables

One can adopt a more abstract point of view. It will be useful for quantum mechanics.

2.1.3.a - Functions as elements of a commutative C^*-algebra

The real functions (continuous, with compact support) on phase space f; $\Omega \to \mathbb{R}$ form a commutative algebra $\mathcal{A}$ with the standard addition and multiplication laws.

$$(f + g)(x) = f(x) + g(x) \quad , \qquad (fg)(x) = f(x)g(x) \qquad (2.1.53)$$

Statistical states (probability distributions on Ω) can be viewed as normalized positive linear forms φ on $\mathcal{A}$, i.e. applications $\mathcal{A} \to \mathbb{R}$ such that

$$\varphi(\alpha f + \beta g) = \alpha \varphi(f) + \beta \varphi(g) \,, \text{with} \quad \alpha, \beta \in \mathbb{R} \tag{2.1.54}$$

$$\varphi(|f|^2) \geq 0 \,, \quad \varphi(1) = 1 \tag{2.1.55}$$

The sup norm or $\mathcal{L}^\infty$ norm, is defined on $\mathcal{A}$ as

$$\|f\|^2 = \sup_{x \in \Omega} |f(x)|^2 = \sup_{\varphi \text{ states}} \varphi(|f(x)|^2) \tag{2.1.56}$$

It has clearly the following properties (extending $\mathcal{A}$ from the algebra of real function to the algebra of complex functions)

$$\|f\| = \|f^*\| \,, \quad \|fg\| \leq \|f\| \, \|g\| \,, \quad \|ff^*\| = \|f\|^2 \tag{2.1.57}$$

and $\mathcal{A}$ is complete under this norm (the limit of a convergent series in $\mathcal{A}$ belongs to $\mathcal{A}$). This makes the algebra $\mathcal{A}$ a mathematical object denoted a commutative C*-algebra.

A famous theorem by Gelfand and Naimark states that the reciprocal is true. Any commutative C*-algebra $\mathcal{A}$ is isomorphic to the algebra $C(X)$ of the continuous functions on some topological (locally compact) space X. This seems a somehow formal result (the space X and its topology may be quite wild, and very far from a regular smooth manifold). What it is important is that a mathematical object (here a topological space X) can be defined intrinsically (by its elements x) or equivalently by the abstract properties of some set of functions over this object (here the commutative algebra of observables). This modern point of view in mathematics (basically this idea is at the basis of the category formalism) is also important in modern physics, as we shall see later in the quantum case.

Technically, the proof is simple. Starting from some X and $\mathcal{A} = C(X)$, to any element $x \in X$, one associate the subalgebra $\mathcal{I}_x$ of all functions that vanish at x

$$\mathcal{I}_x : \{f \in \mathcal{A}; \ f(x) = 0\} \tag{2.1.58}$$

The $\mathcal{I}_x$ are the maximal ideals of $\mathcal{A}$, (left-)ideals $\mathcal{I}$ of an algebra $\mathcal{A}$ being subalgebras of $\mathcal{A}$ such that $x \in \mathcal{I}$ and $y \in \mathcal{A}$ implies $xy \in \mathcal{I}$. It is easy to show that the set $\mathcal{X}$ of the maximal ideals of $\mathcal{A} = C(X)$ is isomorphic to X, and that $\mathcal{A}/\mathcal{I}_x = \mathbb{C}$ the target space. reciprocally, to any commutative C^* algebra $\mathcal{A}$ one can associate the set of its maximal ideas $\mathcal{X}$, show that it is locally compact and that

$$\mathcal{X} = \{\text{maximal ideals of } \mathcal{A}\} \iff \mathcal{A} = C(\mathcal{X}) \tag{2.1.59}$$

2.1.3.b - Observable as elements of a commutative Poisson algebra

For the Hamiltonian systems, the algebra of observables $\mathcal{A}$ of (now C^∞, i.e. differentiable) functions on the phase space Ω is equipped with an additional product, the antisymmetric Poisson bracket $\{\cdot, \cdot\}$ that satisfy 2.1.28, 2.1.29 and 2.1.30. This algebra $\mathcal{A}$, with its three laws (addition, multiplication, Poisson bracket) is now a commutative Poisson algebra. Poisson algebras may be non-commutative.

The most general formulation for classical Hamiltonian dynamics is that of Poisson manifolds and of its associated commutative Poisson algebra. Poisson manifolds are manifolds (phase space) embodied with a Poisson bracket, but this is a more general formulation that symplectic manifolds, since it encompasses special situations where the Poisson bracket is degenerate, and does not define a symplectic structure. They are useful for some general situation, in particular when studying the effective classical dynamics that emerge in some semiclassical limit from a fully quantum dynamics. Poisson manifolds can in general be split (foliated) into "symplectic leaves" embodied with a well defined induced symplectic structure.

The fact that in classical mechanics dynamics are given by Hamiltonian flows on a phase space which is a symplectic or a Poisson manifold can be somehow justified at the classical level. One has to assume that the possible dynamics are given by flows equations generated by some smooth vector fields, that these flows are generated by conserved quantities (Hamiltonians) and that the dynamics are covariant under change of frames generated by these flows (existence and invariance of canonical transformations).

However a full understanding and justification of classical Hamiltonian dynamics comes from quantum mechanics. Indeed, the Poisson bracket structure is the "classical limit" of the Lie algebra structure of the commutators of quantum observables (operators) in quantum mechanics, and the canonical transformations are the classical version of the unitary transformations in the Hilbert space of pure states.

2.2   Probabilities

Probabilities are an important aspect of classical physics [1] and are one of the key components of quantum physics, since the later is intrinsically probabilistic. Without going into any details and much formalism, I think it is important to recall the two main ways to consider and use probabilities in mathematics, statistics, physics and natural sciences: the frequentist point of view and the Bayesian point of view. At the level considered here, these are different point of views on the same mathematical formalism, and on its use. As we shall see, in some sense quantum mechanics forces us to treat them on the same footing. There are of course many different, more subtle and more precise mathematical as well as philosophical points of view on probability theory. I shall not enter in any discussion about the merits and the consistency of objective probabilities versus subjective probabilities.

Amongst many standard references on the mathematical formalism of probability, there is the book by Kolmogorov [Kol50], and the book by Feller [Fel68]. See also the quick introduction for and by a physicist by M. Bauer (in french) [Bau09]. References on Bayesian probabilities are the books by de Finetti [dF74], by Jaynes [Jay03] and the article by Cox [Cox46].

---

1. Probability theory appeared and developed in parallel with classical physics, with important contributors in both fields, from Pascal, Bernouilli, and Laplace to Poincaré and Kolmogorov

2.2.1   The frequentist point of view

The frequentist point of view is the most familiar and the most used in statistical physics, dynamical systems, as well as in mathematics (it is at the basis of the formulation of modern probability theory from the beginning of 20th century, in particular for the Kolmogorov axiomatic formulation of probabilities). Roughly speaking, probabilities represent a measure of ignorance on the state of a system, coming for instance from: uncertainty on its initial state, uncertainty on its dynamical evolution due to uncertainty on the dynamics, or high sensibility to the initial conditions (chaos). Then probabilities are asymptotic frequencies of events (measurements) if we repeat observations on systems prepared by the same initial procedure. More precisely, one has a set Ω of samples (the sample space), a σ-algebra $\mathcal{F}$ of "measurable" subsets of the sample space Ω, and a measure P on $\mathcal{F}$ (equivalently a probability measure μ on Ω). This probability measure is a priori given. Thus to a subset $E \in \mathcal{F}$ is associated an event E and the probability for this event to happen (its expectation)

$$P(E) = \mathbb{E}[x \in E] = \int_\Omega \mu(dx)\chi_E(x) \tag{2.2.1}$$

2.2.2   The Bayesian point of view

The so called Bayesian point of view is somehow broader, and of use in statistics, game theory, economy, but also in experimental sciences. It is also closer to the initial formulations of probabilities (or "chance") in the 18th and 19th centuries. It has been reviewed by statisticians like de Finetti or Jaynes (among others) in the 20th century.

Probabilities are considered as qualitative estimates for the "plausibility" of some proposition (it can be the result of some observation), given some "state of knowledge" on a system.

$$P(A|C) = \text{plausibility of } A, \text{knowing } C \tag{2.2.2}$$

In particular, one considers the "priors"

$$P(C) = P(|C) \tag{2.2.3}$$

These probabilities (degree of plausibility) must satisfy rules that are constrained by logical principles, and which turns out to be the rules of probability theory. This is the so called objectivist point of view (objective probabilities), where the degree of plausibility must be established by a "rational agent" from its knowledge of the system. Another (more controversial) point of view is the "subjectivist point of view" (subjective probabilities) where the probabilities $P(A)$ correspond simply to the "degree of personal belief" of the propositions by the agent. In the former usually the priors are constrained by (for instance) some a priori assumed symmetry principle. The difference between the objective and subjective probabilities will be what are the allowed rules for initial choices for the priors. Note also that the concept of Bayesian probabilities is not accepted by everybody, and may be mathematically problematic when dealing with continuous probability measures.

2.2.3 Conditional probabilities

The basic rules are the same in the different formulations. A most important concept is conditional probabilities $P(A|B)$ (the probability of A, B being given), and the Bayes relation for the conditional probabilities

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{2.2.4}$$

where $P(A)$ and $P(B)$ are the initial probabilities for A and B (the priors), and $P(A|B)$ and $P(B|A)$ the conditional probabilities.

Frequentist: In the frequentist formulation $P(A|B)$ is the frequency of A, once we have selected the samples such that B is true. Bayes formula has the simple representation with Venn diagrams in the set of samples
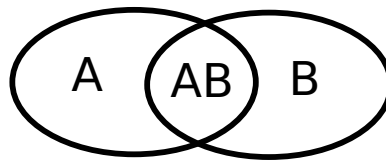

Figure 2.4: Venn representation of the conditional probabilities formula

Bayesian: In the Bayesian formulation (see for instance the book by Jaynes), and has already presented above, one may consider every probabilities as conditional probabilities. For instance $P_C(A) = P(A|C)$, where the proposition C corresponds to the "prior knowledge" that leads to the probability assignment $p_C(A)$ for A (so P_C is the probability distribution). If AB means the proposition "A and B" ($A \wedge B$ or $A + B$), Bayes formula follows from the "product rule"

$$P(AB|C) = P(A|BC)P(B|C) = P(B|AC)P(A|C) \tag{2.2.5}$$

Its meaning is the following: given C, if one already knows the plausibility for AB being true ($P(AB|C)$), and the plausibility for B being true (the prior $P(B|C)$), then 2.2.5 tells us how one should modify the plausibility for A of being true, if we learn that B is moreover true ($P(A|BC)$). Together with the "sum rule"

$$P(A|C) + P(\neg A|C) = 1 \tag{2.2.6}$$

($\neg A$ is the negation of A), these are the basic rules of probability theory in this framework.

2.3 Quantum mechanics: the "canonical formulation"

Let me now recall the so called "canonical formalism" of quantum mechanics. By "standard formalism" I mean nothing but the typical or standard presentation of the

formalism, as it is given (with of course many variants) in most textbooks, prior to (or without) the discussions on the significance and the possible interpretations of the formalism. It relies on the formalism of Hilbert spaces, vector states and wave functions, on the "correspondence principle" when discussing the quantization of a classical (usually non-relativistic) system, and aims of course at physical and calculational understanding and efficiency [2] when discussing physical systems and experiments.

There is of course an enormous number of good books on quantum mechanics and quantum field theory. Among the very first books on quantum mechanics, those of P. A. Dirac (1930) [Dir30] and J. von Neumann (1932) [vN32] ([vN55] for the english traduction of 1955) are still very useful and valuable. Some very good modern books on quantum mechanics with a contemporary view and a treatment of the recent developments are the latest edition of the standard by Cohen-Tanoudji, Diu & Laloe [CTDL06], the book by M. le Bellac [LB11], and the book by Auletta, Fortunato and Parisi [AFP09]. Let me also quote at that stage the book by A. Peres [Per95], although it is much more focused on the conceptual aspects.

Some standard modern references on quantum field theory are the books by J. Zinn-Justin [ZJ02], by S. Weinberg [Wei05] and the book by A. Zee [Zee03] (in a very different relaxed style). Reference more oriented towards mathematical physics will be given later.

Amongst the numerous book and articles on the questions of the foundation and the interpretation of quantum mechanics, very good references are the encyclopedic and balanced review by Auletta [Aul01], and the more recent and shorter book by F. Laloe [Lal12] (see also [Lal11, Lal01]). More will be given later.

2.3.1 Principles

I give here one of the variants of the canonical formalism, without any justification.

2.3.1.a - Pure states and Hilbert space

The phase space Ω of classical mechanics is replaced by the complex Hilbert space $\mathcal{H}$ of pure states. Elements of $\mathcal{H}$ (vectors) are denoted ψ or $|\psi\rangle$ ("kets" in Dirac notations). The scalar product of two vectors ψ and ψ' in $\mathcal{H}$ is denoted $\psi^* \cdot \psi'$ or $\langle\psi|\psi'\rangle$. The $\psi^* = \langle\psi|$ are the "bra" and belong to the dual $\mathcal{H}^*$ of $\mathcal{H}$. Note that in the mathematical litterature the scalar product is often noted in the opposite order $\langle\psi|\psi'\rangle = \psi' \cdot \psi^*$. We shall stick to the physicists notations.

Pure quantum states are rays of the Hilbert space, i.e. 1 dimensional subspaces of $\mathcal{H}$. They correspond to unit norm vectors $|\psi\rangle$, such that $\|\psi\|^2 = \langle\psi|\psi\rangle = 1$, and modulo an arbitrary unphysical (unobservable) phase $|\psi\rangle \simeq e^{i\theta}|\psi\rangle$. This normalization condition comes of course from the Born rule (see below).

Of course a consequence of this principle is that that any complex linear combination of two states $|\alpha\rangle$ and $|\beta\rangle$, $|\psi\rangle = a|\alpha\rangle + b|\beta\rangle$, corresponds also [3] a state of the system. This is the so-called superposition principle.

---

2. Looking for efficiency and operability does not mean adopting the (in)famous "shut up and calculate" stance, an advice often but falsely attributed to R. Feynman

3. At least for finite dimensional and simple cases of infinite dimensional Hilbert spaces, see the discussion on superselection sectors

The complex structure of the Hilbert space is a crucial feature of quantum mechanics. It is reflected in classical mechanics by the almost complex and symplectic structures of the classical phase space.

2.3.1.b - Observables and operators

The physical observables A are the self-adjoint operators on $\mathcal{H}$ (Hermitian or symmetric operators), such that $A = A^\dagger$, where the conjugation is defined by $\langle A^\dagger \psi' | \psi \rangle = \langle \psi' | A\psi \rangle$. Note that the conjugation $A^\dagger$ is rather denoted A^* in the mathematical literature, and in some sections we shall use this notation, when dealing with general Hilbert spaces not necessarily complex.

The operators on $\mathcal{H}$ form an associative, but non commutative complex operator algebra. Any set of commuting self-adjoint operators $\{A_i\}$ corresponds to a family of classically compatible observables, which in principle can be measured independently.

2.3.1.c - Measurements, probabilities and the Born rule

The outcome of the measurement of an observable A on the system in a state ψ is in general not deterministic. Quantum mechanics give only probabilities for the outcomes, and in particular the expectation value of the outcomes $\langle A \rangle_\psi$. This expectation value is given by the Born rule

$$\langle A \rangle_\psi = \langle \psi | A | \psi \rangle = \langle \psi | A\psi \rangle \tag{2.3.1}$$

For compatible (commuting) observables the probabilities of outcome obey the standard rule of probabilities and these measurements can be repeated and performed independently.

This implies (or is equivalent to state) that the possible outcomes of the measurement of A must belong to the spectrum of A, i.e. can only equal the eigenvalues of A (I consider the simple case where A has a discrete spectrum). Moreover the probability $p_i(\psi)$ to get as outcome the eigenvalue a_i, denoting $|i\rangle$ the corresponding eigenvector, is the modulus squared of the probability amplitude $\langle i | \psi \rangle$

$$p_i(\psi) \;=\; \begin{array}{c} \text{probability of outcome of } A \to a_i \\ \text{if the system is in the state } |\psi\rangle \end{array} \;=\; |\langle i | \psi \rangle|^2$$

A very important consequence (or feature) is that quantum measurements are intrinsically irreversible processes. Let us consider ideal measurements (non destructive measurements), i.e. measurement operations which can be repeated quasi-instantaneously on quantum systems and when repeated, give always the same result. If the system is initially in a state $\psi\rangle$, if one performs an ideal measurement of A with outcome a_i, and if there is single eigenvector $|i\rangle$ associated to this eigenvalue a_i of A, this implies that the system must be considered to be in the associated eigenstate $|i\rangle$. If the eigenspace V_i associated to the eigenvalue a_i is a higher dimensional subspace $V_i \subset \mathcal{H}$, then the system must be in the projected state $|\psi_i\rangle = P_i|\psi\rangle$, with P_i the orthogonal projector onto V_i. This is the projection postulate.

For observables with a continuous spectrum and non-normalizable eignestates, the theory of ideal measurements involves more mathematical rigor and the use of spectral theory.

At that stage I do not discuss what it means to "prepare a system in a given state", what "represents" the state vector, what is really a measurement process (the problem of quantum measurement) and what means the projection postulate. We shall come back to some of these questions along the course.

2.3.1.d - Unitary dynamics

For a closed system, the time evolution of the states is linear and it must preserve the probabilities, hence the scalar product $\langle .|. \rangle$. Therefore is given by unitary transformations $U(t)$ such that $U^{-1} = U^{\dagger}$. Again if the system is isolated the time evolution form a multiplicative group acting on the Hilbert space and its algebra of observables, hence it is generated by an Hamiltonian self-adjoint operator H

$$U(t) = \exp\left(\frac{t}{i\hbar} H\right)$$

The evolution equations for states and observables are discussed below.

2.3.1.e - Multipartite systems

Assuming that it is possible to perform independent measurements on two independent (causally) subsystems S_1 and S_2 implies (at least in the finite dimensional case) that the Hilbert space $\mathcal{H}$ of the states of the composite system $S = "S_1 \cup S_2"$ is the tensor product of the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ of the two subsystems.

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$$

This implies the existence for the system $\mathcal{S}$ of generic "entangled states" between the two subsystems

$$|\Psi\rangle = c|\psi\rangle_1 \otimes |\phi\rangle_2 + c'|\psi'\rangle_1 \otimes |\phi'\rangle_2$$

Entanglement is one of the most important feature of quantum mechanics, and has no counterpart in classical mechanics. It is entanglement that leads to many of the counter-intuitive features of quantum mechanics, but it leads also to many of its interesting aspects and to some of its greatest successes.

Finally let me insist on this point: the fact that one can treat parts of a physical system as independent subparts is not obvious. One may ask what is the property that characterizes this fact that two systems are independent, in the sense that they are causally independent. In fact two parts of a system can be considered as independent if all the physical observables relative to the first part commutes with those relative to the second one. This can be properly understood only in the framework of special relativity, and is deeply related to the concept of locality and causality in relativistic quantum theories, namely in quantum field theories.

2.3.1.f - Correspondence principe, canonical quantization

The correspondence principle has been very important in the elaboration of quantum mechanics. Here by correspondence principle I mean that when quantizing a classical system, often one can associate to canonically conjugate variables (q_i, p_i) self-adjoint operators (Q_i, P_i) that satisfy the canonical commutation relations

$$\{q_i, p_i\} = \delta_{ij} \implies [Q_i, P_j] = i\hbar \delta_{ij} \tag{2.3.2}$$

and to take as Hamiltonian the operator obtained by replacing in the classical Hamiltonian the variables (q_i, p_i) by the corresponding operators.

For instance, for the particle on a line in a potential, one takes as (Q, P) the position and the momentum and for the Hamiltonian

$$H = \frac{P^2}{2m} + V(Q) \tag{2.3.3}$$

The usual explicit representation is, starting from the classical position space $\mathbb{R}$ as configuration space, to take for Hilbert space the space of square integrable functions $\mathcal{H} = \mathcal{L}^2(\mathbb{R})$, the states $|\psi\rangle$ correspond to the wave functions $\psi(q)$, and the operators are represented as

$$Q = q \quad , \qquad P = \frac{\hbar}{i} \frac{\partial}{\partial q} \tag{2.3.4}$$

Of course the value of the wave function $\psi(q)$ is simply the scalar product of the state $|\psi\rangle$ with the position eigenstate $|q\rangle$

$$\psi(q) = \langle q|\psi\rangle \quad , \qquad Q|q\rangle = q|q\rangle \tag{2.3.5}$$

(a proper mathematical formulation involving the formalism of ringed Hilbert spaces).

2.3.2   Representations of quantum mechanics

The representation of states and observables as vectors and operators is invariant under global unitary transformations (the analog of canonical transformations in classical mechanics). These unitary transformations may depend on time. Therefore there are different representations of the dynamics in quantum mechanics. I recall the two main representations.

2.3.2.a - The Schrödinger picture

It is the most simple, and the most used in non relativistic quantum mechanics, in canonical quantization and is useful to formulate the path integral. In the Schrödinger picture the states ψ (the kets $|\psi\rangle$) evolve with time and are noted $\psi(t)$. The observables are represented by time independent operators. The evolution is given by the Schrödinger equation

$$i\hbar \frac{d\psi}{dt} = H\psi \tag{2.3.6}$$

The expectation value of an observable A measured at time t for a system in the state ψ is thus

$$\langle A \rangle_{\psi(t)} = \langle \psi(t)|A|\psi(t)\rangle \tag{2.3.7}$$

The evolution operator $U(t)$ is defined by

$$\psi(t=0) = \psi_0 \quad \to \quad \psi(t) = U(t)\psi_0 \tag{2.3.8}$$

It is given by

$$U(t) = \exp\left(\frac{t}{i\hbar}H\right) \tag{2.3.9}$$

and obeys the evolution equation

$$i\hbar\frac{d}{dt}U(t) = H\,U(t) \qquad ; \qquad U(0) = \mathbf{1} \tag{2.3.10}$$

This generalizes easily to the case where the Hamiltonian depends explicitly of the time t. Then

$$i\hbar\frac{d}{dt}U(t,t_0) = H(t)\,U(t,t_0) \qquad ; \qquad U(t_0,t_0) = \mathbf{1} \tag{2.3.11}$$

and

$$U(t,t_0) = T\left[\exp\left(\frac{1}{i\hbar}\int_{t_0}^{t}dt\,H(t)\right)\right] = \sum_{k=0}^{\infty}(i\hbar)^{-k}\int_{t_0<t_1<\cdots<t_k<t}dt_1\cdots dt_k\,H(t_k)\cdots H(t_1) \tag{2.3.12}$$

where T means the time ordered product (more later).

2.3.2.b - The Heisenberg picture

This representation is the most useful in relativistic quantum field theory. It is in fact the best mathematically fully consistent formulation, since the notion of state in more subtle, in particular it depends on the reference frame. It is required for building the relation between critical systems and Euclidean quantum field theory (statistical field theory).

In the Heisenberg representation, the states are redefined as a function of time via the unitary transformation $U(-t)$ on $\mathcal{H}$, where $U(t)$ is the evolution operator for the Hamiltonian H. They are denoted

$$|\psi;t\rangle = U(-t)|\psi\rangle \tag{2.3.13}$$

The unitary transformation redefines the observables A. They becomes time dependent and are denoted $A(t)$

$$A(t) = U(-t)AU(t) \tag{2.3.14}$$

The dynamics given by the Schrödinger equation is reabsorbed by the unitary transformation. The dynamical states are independent of time!

$$|\psi(t);t\rangle = U(-t)U(t)|\psi\rangle = |\psi\rangle \tag{2.3.15}$$

The expectation value of an observable A on a state ψ at time t is in the Heisenberg representation

$$\langle A(t)\rangle_\psi = \langle\psi(t);t|A(t)|\psi(t;,t\rangle = \langle\psi|A(t)|\psi\rangle \tag{2.3.16}$$

The Schrödinger and Heisenberg representation are indeed equivalent, since they give
the same result for the physical observable (the expectation values)

$$\langle A \rangle_{\psi(t)} = \langle A(t) \rangle_{\psi} \tag{2.3.17}$$

In the Heisenberg representation the Hamitonian H remains independent of time
(since it commutes with $U(t)$

$$H(t) = H \tag{2.3.18}$$

The time evolution of the operators is given by the evolution equation

$$i\hbar \frac{d}{dt} A(t) = [A(t), H] \tag{2.3.19}$$

This is the quantum version of the classical Liouville equation 2.1.52. Of course
the Schrödinger and the Heisenberg representations are the quantum analog of the
two "Eulerian" and' "Lagrangian" representations of classical mechanics discussed
above.

For the particle in a potential the equations for Q and P are the quantum version
of the classical Hamilton equations of motion

$$\frac{d}{dt} Q(t) = \frac{1}{m} P(t) \qquad , \qquad \frac{d}{dt} P(t) = -V'(Q(t)) \tag{2.3.20}$$

For an observable A which depends explicitly of time (in the Schrödinger picture), the
evolution equation becomes

$$i\hbar \frac{d}{dt} A(t) = i\hbar \frac{\partial}{\partial t} A(t) + [A(t), H] \tag{2.3.21}$$

and taking its expectation value in some state ψ one obtains Ehrenfest theorem

$$i\hbar \frac{d}{dt} \langle A \rangle(t) = i\hbar \frac{\partial}{\partial t} \langle A \rangle(t) + \langle [A, H] \rangle(t) \tag{2.3.22}$$

2.3.3 Quantum statistics and the density matrix

2.3.3.a - The density matrix

As in classical physics, in general only some partial information is available on the
physical system one is interested in, or one wants to consider statistics over ensem-
bles of states. Such situations have to be described by the concept of statistical or
mixed state. But in quantum mechanics all the information one can get on a system is
provided by the expectation values of its observables, since quantum mechanics con-
tains some intrinsic indeterminism, and involves already probabilities and statistics.
The pure quantum states $|\psi\rangle$ considered up to now are the quantum states which have
the property that a maximal amount of information can be extracted by appropriate
sets of compatible measurements on the state. The difference with classical physics
is that different maximal sets of information can be extracted from the same state if
one chose to perform different incompatible sets of measurements.

The mathematical concept that represents a general mixed state is the concept of density matrix. But before discussing this, one can start by noticing that, as in classical physics, an abstract statistical state ω is fully characterized by the ensemble of the expectation values $\langle A \rangle_\omega$ of all the observables A of the system, measured over the state ω.

$$\langle \mathbf{A} \rangle_\omega \quad = \text{ expectation value of } \mathbf{A} \text{ measured over the state } \omega \qquad (2.3.23)$$

I denote general statistical states by Greek letters (here ω) and pure states by the bra-ket notation when there is a ambiguity. The ω here should not be confused for the notation for the symplectic form over the classical phase space of a classical system. We are dealing with quantum systems and there is no classical phase space anymore.

From the fact that the observables may be represented as an algebra $\mathcal{A}$ of operators over the Hilbert space $\mathcal{H}$, it is natural to consider that statistical states ω corresponds to linear forms over the algebra of operators $\mathcal{A}$, hence applications $\mathcal{A} \to \mathbb{C}$; $\mathbf{A} \to \langle \mathbf{A} \rangle_\omega$, with the properties

$$\langle a\mathbf{A} + b\mathbf{B} \rangle_\omega = a\langle \mathbf{A} \rangle_\omega + b\langle \mathbf{B} \rangle_\omega \qquad \text{linearity} \qquad (2.3.24)$$

$$\langle \mathbf{A}^\dagger \rangle_\omega = \overline{\langle \mathbf{A} \rangle_\omega} \qquad \text{reality, } \bar{z} \text{ means the complex conjugate of } z \in \mathbb{C} \qquad (2.3.25)$$

$$\langle \mathbf{A}^\dagger \mathbf{A} \rangle_\omega \geq 0 \qquad \text{and} \qquad \langle \mathbf{1} \rangle_\omega = 1 \qquad \text{positivity and normalization} \qquad (2.3.26)$$

For finite dimensional Hilbert spaces and for the most common infinite dimensional cases (for physicists), any such liner a form can be represented as a normalized positive self-adjoint matrix ρ_ω

$$\rho_\omega \geq 0 \quad , \qquad \text{tr}(\rho_\omega) = 1 \qquad (2.3.27)$$

such that for any operator $A \in \mathcal{A}$, its expectation value in the state ω is given by

$$\langle \mathbf{A} \rangle_\omega = \text{tr}(\rho_\omega \mathbf{A}) \qquad (2.3.28)$$

ρ_ω is the density matrix or density operator associated to the state ω. The concept of density matrix was introduced by J. von Neumann (and independently by L. Landau and F. Bloch) in 1927. The identity 2.3.32 is simply the generalization of the Born rule for statistical states.

For pure states $\omega = |\psi\rangle$ the density operator is simply the rank 1 projection operator onto the state $|\psi\rangle$

$$\rho_\psi = |\psi\rangle\langle\psi| \qquad (2.3.29)$$

One can also remark that the set of mixed states, as represented by density matrices, form a convex set (the set of matrices satisfying 2.3.27). It is convex since any statistical mixture of two mixed states ρ_1 and ρ_2 is a mixed state $\rho = p_1\rho_1 + p_2\rho_2$ ($p_1, p_2 \geq 0$ and $p_1 + p_2 = 1$). Pure states are nothing but the extremal points of the set of mixed states, i.e. the states that cannot be written as a mixture of two different states $\rho = p_1\rho_1 + p_2\rho_2$ ($\rho_1 \neq \rho_2$, $p_1, p_2 > 0$ and $p_1 + p_2 = 1$).

Before discussing some properties and features of the density matrix, let me just mention that in the physics literature, the term "state" is usually reserved to pure states, while in the mathematics literature the term "state" is used for general statistical states. The denomination "pure state" or "extremal state" is used for vectors in the Hilbert state and the associated projector. There are in fact some good mathematical reasons to use this general denomination of state.

2.3.3.b - Quantum ensembles versus classical ensembles

Let us consider a system whose Hilbert space is finite dimensional ($\dim(\mathcal{H}) = N$), in a state given by a density matrix ρ_ω. ρ_ω is a $N \times N$ self-adjoint positive matrix. It is diagonalizable and its eigenvalues are ≥ 0. If it has $1 \leq K \leq N$ orthonormal eigenvectors labeled by $|n\rangle$ ($n = 1, \cdots K$) associated with K non-zero eigenvalues p_n ($n = 1, \cdots K$) one can write

$$\rho_\omega = \sum_{n=1}^{K} p_n |n\rangle\langle n| \tag{2.3.30}$$

with

$$0 < p_n \leq 1 , \quad \sum_n p_n = 1 \tag{2.3.31}$$

The expectation value of any observable A in the state ω is

$$\langle \mathsf{A} \rangle_\omega = \sum_n p_n \langle n|\mathsf{A}|n\rangle \tag{2.3.32}$$

The statistical state ω can therefore be viewed as a classical statistical mixture of the K orthonormal pure states $|n\rangle$, $n = 1, \cdots K$, the probability of the system to be in the pure state $|n\rangle$ being equal to p_n.

This point of view is useful but may be misleading. It should not be used to infer statements on how the system has been prepared. One can indeed build a statistical ensemble of independently prepared copies of the system corresponding to the state ω by picking at random, with probability p_n the system in the state $|n\rangle$. But this is not the only way to build a statistical ensemble corresponding to ω. More precisely, there are many different ways to prepare a statistical ensemble of states for the system, by picking with some probability p_α copies of the system in different states among a pre chosen set $\{|\psi_\alpha\rangle\}$ of (a priori not necessarily orthonormal) pure states, which give the same density matrix ρ_ω.

This is not a paradox. The difference between the different preparation modes is contained in the quantum correlations between the (copies of the) system and the devices used to do the preparation. These quantum correlations are fully inaccessible if one performs measurements on the system alone. The density matrix contains only the information about the statistics of the measurement on the system alone (but it encodes the maximally available information obtainable by measurements on the system only).

Another subtle point is that an ensemble of copies of a system is described by a density matrix ρ for the single system if the different copies are really independent, i.e. if there are no correlations between different copies in the ensemble, or if one simply neglect (project out) these correlations. Some apparent paradoxes arise if there are such correlations and if they must be taken into account. One must then consider the matrix density for several copies, taken as a larger composite quantum system.

2.3.3.c - The von Neumann entropy

The "degree of uncertainty" or "lack of information" which is "contained in" a mixed quantum state ω is given by the von Neumann entropy

$$S(\omega) = -\operatorname{tr}(\rho_\omega \log\rho_\omega) = -\sum_n p_n \log p_n \qquad (2.3.33)$$

It is the analog of the Boltzman-Gibbs entropy for a classical statistical distribution. It shares also some deep relation with Shannon entropy in information theory (more later).

The entropy of a pure state is minimal and zero. Conversely, the state of maximal entropy is the statistical state where all quantum pure states are equiprobable. It is given by a density matrix proportional to the identity, and the entropy is the logarithm of the number of accessible different (orthogonal) pure quantum state, i.e. of the dimension of the Hilbert space (in agreement with the famous Boltzmann formula $W = k_B \log N$).

$$\rho = \frac{1}{N}\mathbf{1} \quad , \quad S = \log N \quad , \qquad N = \dim\mathcal{H} \qquad (2.3.34)$$

2.3.3.d - Example: entanglement entropy

An important context where the density matrix has to be used is the context of open quantum systems and multipartite quantum systems. Consider a bipartite system $\mathcal{S}$ composed of two distinct subsystems $\mathcal{A}$ and $\mathcal{B}$. The Hilbert space $\mathcal{H}_\mathcal{S}$ of the pure states of $\mathcal{S}$ is the tensor product of the Hilbert space of the two subsystems

$$\mathcal{H}_\mathcal{S} = \mathcal{H}_\mathcal{A} \otimes \mathcal{H}_\mathcal{B} \qquad (2.3.35)$$

Let us assume that the total system is in a statistical state given by a density matrix $\rho_\mathcal{S}$, but that one is interested only in the subsystem $\mathcal{A}$ (or $\mathcal{B}$). In particular one can only perform easement on observables relative to $\mathcal{A}$ (or $\mathcal{B}$). Then all the information on $\mathcal{A}$ is contained in the reduced density matrix $\rho_\mathcal{A}$; obtained by taking the partial trace of the density matrix for the whole system $\rho_\mathcal{S}$ over the (matrix indices relative to the) system $\mathcal{B}$.

$$\rho_\mathcal{A} = \operatorname{tr}_\mathcal{B}[\rho_\mathcal{S}] \qquad (2.3.36)$$

This is simply the quantum analog of taking the marginal of a probability distribution $p(x, y)$ with respect to one of the random variables $\rho_x(x) = \int dy\, \rho(x, y)$).

If the system $\mathcal{S}$ is in a pure state $|\psi\rangle$, but if this state is entangled between $\mathcal{A}$ and $\mathcal{B}$, the reduced density matrix $\rho_\mathcal{A}$ is that of a mixed state, and its entropy is $S_A(\rho_\mathcal{A}) > 0$. Indeed when considering $\mathcal{A}$ only the quantum correlations between $\mathcal{A}$ and $\mathcal{B}$ have been lost. If $\mathcal{S}$ is in a pure state the entropies $S_A(\rho_\mathcal{A}) = S_B(\rho_\mathcal{B})$. This entropy is then called the entanglement entropy. Let us just recall that this is precisely one of the context where the concept of von Neumann entropy was introduced around 1927. More properties of features of quantum entropies will be given later.

2.3.3.e - Thermal states

A standard example of density matrix is provided by considering an quantum system $\mathcal{S}$ which is (weakly) coupled to a large thermostat, so that it is at equilibrium, exchanging freely energy (as well as other quantum correlations) with the thermostat, and at a finite temperature T. Then the mixed state of the system is a thermal Gibbs state (or in full generality a Kubo-Martin-Schwinger or KMS state). If the spectrum of the Hamiltonian H of the system is discrete, with the eigenstates $|n\rangle$, $n \in \mathbb{N}$ and eigenvalues (energy levels) by E_n (with $E_0 < E_1 < E_2 \cdots$), the density matrix is

$$\rho_\beta = \frac{1}{Z(\beta)} \exp(-\beta H) \tag{2.3.37}$$

with $Z(\beta)$ the partition function

$$Z(\beta) = \mathrm{tr}\left[\exp(-\beta H)\right] \tag{2.3.38}$$

and

$$\beta = \frac{1}{k_B T} \tag{2.3.39}$$

In the energy eigenstates basis the density matrix reads

$$\rho_\beta = \sum_n p_n |n\rangle\langle n| \tag{2.3.40}$$

with p_n the standard Gibbs probability

$$p_n = \frac{1}{Z(\beta)} \exp(-\beta E_n) \qquad ; \qquad Z(\beta) = \sum_n \exp(-\beta E_n) \tag{2.3.41}$$

The expectation value of an observable A in the thermal state at temperature T is

$$\langle A \rangle_\beta = \sum_n p_n \langle n|A|n\rangle = \frac{\mathrm{tr}\left[A \exp(-\beta H)\right]}{\mathrm{tr}\left[\exp(-\beta H)\right]} \tag{2.3.42}$$

For infinite systems with an infinite number of degrees of freedom, several equilibrium macroscopic states may coexist. The density matrix formalism is not sufficient and must be replaced by the formalism of KMS states (Kubo-Martin-Schwinger). This will be discussed a bit more later in connection with superselection sectors in the algebraic formalism.

2.3.3.f - Imaginary time formalism

Let us come back to the simple case of a quantum non-relativistic system, whose energy spectrum is bounded below (and discrete to make things simple), but unbounded from above. The evolution operator

$$U(t) = \exp\left(\frac{t}{i\hbar} H\right) \tag{2.3.43}$$

considered as a function of the time t, may be extended from "physical" real time $t \in \mathbb{R}$ to complex time variable, provided that

$$\text{Im}(t) \leq 0 \tag{2.3.44}$$

More precisely, $U(t)$ as an operator, belongs to the algebra $\mathcal{B}(\mathcal{H})$ of bounded operators on the Hilbert space $\mathcal{H}$. A bounded operator A on $\mathcal{H}$ is an operator whose L^∞ norm, defined as

$$\|A\|^2 = \sup_{\psi \in \mathcal{H}} \frac{\langle \psi | A^\dagger A | \psi \rangle}{\langle \psi | \psi \rangle} \tag{2.3.45}$$

is finite. This is clear in the simple case where

$$U(t) = \sum_n \exp\left(\frac{t}{i\hbar} E_n\right) |n\rangle\langle n| \quad , \quad \|U(t)\| = \begin{cases} \exp\left(\frac{\text{Im}(t)}{\hbar} E_0\right) & \text{if } \text{Im}(t) \leq 0, \\ +\infty & \text{otherwise.} \end{cases} \tag{2.3.46}$$

The properties of the algebras of bounded operators and of their norm will be discussed in more details in the next section on the algebraic formulation of quantum mechanics.
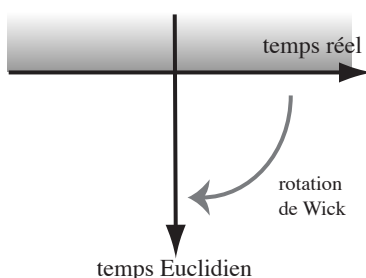

Figure 2.5: Real time t and imaginary (Euclidean) time $\tau = i t$: Wick rotation

Consider now the case where t is purely imaginary

$$t = -i\tau \quad , \quad \tau > 0 \quad \text{real} \qquad U(-i\tau) = \exp\left(-\frac{\tau}{\hbar} H\right) \tag{2.3.47}$$

The evolution operator has the same form than the density matrix for the system in a Gibbs state at temperature T

$$\rho_\beta = \frac{1}{Z(\beta)} U(-i\tau) \quad , \quad \beta = \frac{1}{k_B T} = \frac{\tau}{\hbar} = i \frac{t}{\hbar} \tag{2.3.48}$$

There is deep analogy

$$\text{imaginary time} \sim \text{finite temperature}$$

Moreover, when considering relativistic quantum field theories in a Lorentzian metric $ds^2 = -dt^2 + d\vec{x}^2$, considering the theory at imaginary time $t = -i\tau$ implies that this

imaginary time τ becomes an "Euclidean coordinate" $\tau = x^0$, and Minkowski space time becomes Euclidean space, with metric $ds^2 = d\tau^2 + d\vec{x}^2$

These seemingly formal analogies are in fact quite important and have numerous applications. They are at the basis of Euclidean Field Theory and of the many applications of quantum field theory to statistical physics, condensed matter and probabilities. Reciprocally, statistical physics methods have found many applications in quantum physics and high energy physics (for instance lattice gauge theories). Considering quantum theory for imaginary time is also very useful in high energy physics, and in quantum gravity. Finally this relation between Gibbs (KMS) states and the unitary evolution operator extends tin mathematics to a more general relation between states and automorphisms of operator algebras (the Tomita-Takesaki theory), that we shall discuss (very superficially) in the next chapter.

2.4 Path and functional integrals formulations

2.4.1 Path integrals

2.4.1.a - Path integral in configuration space

It is known since Feynman that a very useful and efficient, if usually not yet mathematically rigorous, way to represent matrix elements of the evolution operator of a quantum system (the transition amplitudes, or "propagators") is provided by path integrals (for non-relativistic systems with a few degrees of freedom) and functional integrals (for relativistic or non relativistic systems with continuous degrees of freedoms, i.e. quantum fields).

Standard references on path integral methods on quantum mechanics and quantum field theory are the original book by Feynman & Hibbs [RPF10], and the books by J. Zinn-Justin [ZJ02], [ZJ10].

For a single particle in an external potential this probability amplitude K for propagation from q_i at time t_i to q_f at time t_f

$$\langle q_f | U(t_f - t_i) | q_i \rangle = \langle q_f, t_f | q_i, t_i \rangle \qquad U(t) = \exp\left(\frac{t}{i\hbar} H\right) \qquad (2.4.1)$$

(the first notation refers to the Schrödinger picture, the second one to the Heisenberg picture) can be written as a sum of histories (or path) $q = \{q(t); t_i \leq t \leq t_f\}$ starting from q_i at time t_i and ending at q_f at time t_f

$$\int_{\substack{q(t_i)=q_I \\ q(t_f)=q_f}} \mathcal{D}[q] \exp\left(\frac{i}{\hbar} S[q]\right) \qquad (2.4.2)$$

where $S[q]$ is the classical action of the trajectory (history) .

The precise derivation of this formula, as well as its proper mathematical definition, is obtained by decomposing the evolution of the system in a large number N of evolutions during elementary time step $\Delta t = \epsilon = t/N$, at arbitrary intermediate positions $q(t_n = n\epsilon)$, $n \in \{1, \cdots, N-1\}$, using the superposition principle. One then uses the
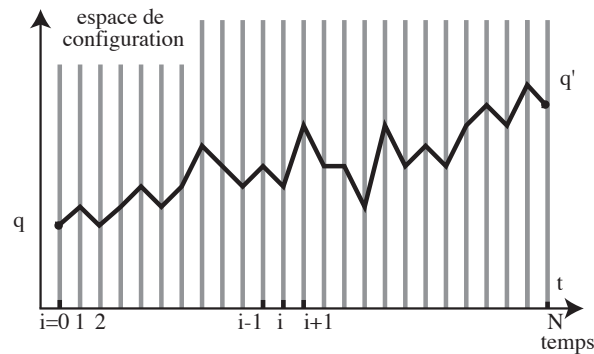
Figure 2.6: Path integral: time discretization

explicit formula for the propagation kernel at small time (the potential $V(q)$ may be considered as constant locally)

$$K(q_f, \epsilon, q_i, 0) \simeq \left(\frac{2i\pi\hbar\epsilon}{m}\right)^{-1/2} \exp\left(\frac{i}{\hbar}\left(\frac{m}{2}\frac{(q_f - q_i)^2}{\epsilon} - \epsilon V\left(\frac{q_f + q_i}{2}\right)\right)\right) \qquad (2.4.3)$$

and one then takes the continuous time limit $\epsilon \to 0$. The precise definition of the measure over histories or paths is (from the prefactor)

$$\mathcal{D}[q] = \prod_{n=1}^{N-1}\left(dq(t_n)\left(\frac{2i\pi\hbar\epsilon}{m}\right)^{-1/2}\right) \qquad (2.4.4)$$

2.4.1.b - Path integral in phase space

This Lagrangian formulation of the path integral rely on a specific choice of configuration space, here the physical space of positions for the single particle. One should keep in mind that different path integrals may correspond to different quantization schemes of the same quantum theory. In particular, for this system the "Lagrangian" path integral has a "Hamiltonian" version, which corresponds to a path integral in phase space. It reads as a sum over trajectories in phase space $\{(q(t), p(t); t_i < t \le t_f\}$

$$\int_{q(t_i)=q_i,\ q(t_f)=q_f} \mathcal{D}[q, p] \exp\left(\frac{i}{\hbar}\int dt\,(p\dot{q} - H(q, p)))\right) \qquad (2.4.5)$$

But one must be very careful on the definition of this path integral (discretization and continuum time limit) and on the definition of the measure $\mathcal{D}[q, p]$ in order to obtain a consistent quantum theory.

2.4.2   Field theories, functional integrals

The path integral representations can be generalized to relativistic quantum field theories. Let us consider the free scalar field, whose classical action (corresponding to the Klein-Gordon equation) is

$$
S[\phi] = \int dt \int d^3\vec{x}\, \frac{1}{2}\left(\left(\frac{\partial\phi}{\partial t}\right)^2 - \left(\frac{\partial\phi}{\partial\vec{x}}\right)^2 - m^2\phi^2\right) = \int d^4x\, \frac{1}{2}\left(-\partial^\mu\phi\,\partial_\mu\phi - m^2\phi^2\right) \quad (2.4.6)
$$

The path integral becomes a functional integral over field configurations ϕ over spacetime $\mathbf{M}^{1,3}$ of the form

$$
\int \mathcal{D}[\phi]\, e^{\frac{i}{\hbar}S[\phi]} \quad , \qquad \mathcal{D}[\phi] \simeq \prod_{x\in\mathbf{M}^{1,3}} d\phi(x) \qquad (2.4.7)
$$

More precisely, the vacuum expectation value of time ordered product of local field operators ϕ in this quantum field theory (the so called Wightman functions, or correlation functions) can be expressed as functional integrals

$$
\langle\Omega|T\phi(x_1)\cdots\phi(x_N)|\Omega\rangle = \frac{1}{Z}\int \mathcal{D}[\phi]\, e^{\frac{i}{\hbar}S[\phi]}\phi(x_1)\cdots\phi(x_N) \qquad (2.4.8)
$$

Z is the partition function or vacuum amplitude

$$
Z = \int \mathcal{D}[\phi]\, e^{\frac{i}{\hbar}S[\phi]} \qquad (2.4.9)
$$

The factor Z is a normalization factor for the functional integral, so that the vacuum to vacuum transition amplitude is

$$
\langle\Omega|\Omega\rangle = 1
$$

The functional integral quantization method requires much care to be defined properly with full mathematical rigor. In particular the high energy-momenta / short distance singularities of quantum fields require the theory of renormalization to construct the functional integral and to check if indeed a continuum relativistic theory can be obtained and make sense as a quantum theory. This is known to be the case in some cases only (QFT in dimensions $D < 4$, some theories in $D = 4$ with an infra-red cut-off, i.e. in a finite volume).

The path integral and functional integral formulations are nevertheless invaluable tools to formulate many quantum systems and quantum field theories, grasp some of their perturbative and non-perturbative features, and perform explicit calculations. They give in particular a very simple and intuitive picture of the semiclassical regimes. they explain why the laws of classical physics can be formulated via variational principles, since classical trajectories are just the stationary phase trajectories (saddle points) dominating the sum over trajectories in the classical limit $\hbar \to 0$. In many cases they allow to treat and visualize quantum interference effects when a few semiclassical trajectories dominates (for instance for trace formulas).

Functional integral methods are also very important conceptually for quantum field theory: from the renormalization of QED to the quantization and proof of renormalisability of non abelian gauge theories, the treatment of topological effects and anomalies in QFT, the formulation of the Wilsonian renormalization group, the applications of QFT methods to statistical mechanics, etc. They thus provides a very useful way to quantize a theory, at least in semiclassical regime where one expect that the quantum theory is not in a too strong coupling regime and where quantum correlations and interference effects (possibly between non-trivial topological sectors) can be kept under control.

I shall not elaborate further here. When discussing the quantum formalism, one should keep in mind that the path integrals and functional integrals represent a very useful and powerful (if usually not fully mathematically rigorous) way to visualize, manipulate and compute transition amplitudes, i.e. matrix elements of operators. Thus functional integrals rather represent an application of the standard canonical formalism, allowing to construct the Hilbert space (or part of it) and the matrix elements of operators of a quantum theory out of the classical theory via a relatively quick and efficient recipe.

2.5 Quantum probabilities and reversibility

2.5.1 Is quantum mechanics reversible or irreversible?

An important aspect of classical physics and of quantum physics is the property of reversibility. By reversibility it is meant that the general formulation of the basic physical laws must be similar under time reversal. This is often stated as:

<p align="center">"There is no microscopic time arrow."</p>

This does not mean that the fundamental interactions (the specific physical laws that govern our universe) are invariant under time reversal. It is known from quantum field theory that (assuming unitarity, locality and Lorentz invariance) any theory must be invariant under CPT only, the product of charge conjugation, parity and time reversal. The reversibility statement means that the dynamics of any given state, viewed forward in time (press key $\boxed{\blacktriangleright}$), is similar to the dynamics, viewed backward in time (press key $\boxed{\blacktriangleleft}$), of some other state (not necessarily the same).

This principle of reversibility is of course also very different from the macroscopic irreversibility that we experience in everyday life (the cosmological arrow of time and the expansion of the universe, the second principle of thermodynamics, some aspects of quantum measurements, irreversible behavious in complex systems such as the Parkinson's laws [Par55] , etc.). I am not going to discuss the issue of the emergence of irreversibility in classical physics, this would require a whole course on dynamical systems and statistical physics. Some elementary aspects will be presented in section 5.5 where quantum measurements are discussed.

In classical mechanics microscopic reversibility is an obvious consequence of the Hamiltonian formulation. In quantum mechanics things are more subtle. Indeed if the evolution of a "closed system" (with no interaction with its environment and the

observer) is deterministic, unitary and reversible (in particular possible quantum cor-
relations between the system and its "outside" are kept untouched), the measurement
processes over quantum systems are known to be irreversible and non-deterministic.
In particular ideal projective measurements feature (via the projection postulate) the
famous phenomenon of "reduction" or 'collapse" of the wave function.

<center>"The measurement process is an irreversible process"</center>

This dichotomy between these two extreme classes of evolution processes has been
known and has been discussed since the birth of quantum mechanics. Is the irre-
versible character of measurement processes a signal of the incomplete character of
quantum mechanics? Is it a strange but unavoidable feature of the quantum word?
Is it a macroscopic effect not so different from the occurrence of irreversibility in the
classical word? Is it related to the cosmological arrow of time or to some quantum
gravity effect?

I am not going to discuss these important questions here. I shall come back to a
few of them in the last part of these notes. Let me just point out that, at the level
of the formalism, the concepts of probabilities and of indeterminism associated with
quantum measurement do not really contradict the principle of microsopic reversibil-
ity, in a certain sense that will be illustrated on a simple example below. This is in fact
known since quite a long time, see for instance the well-known '64 paper by Aharonov,
Bergmann & Lebowitz [ABL64]. Since this point will be very important in this presen-
tation, especially when discussing the quantum logic formalism, let me explain it on
a simple, but basic example, with the usual suspects involved in quantum measure-
ments.

2.5.2   Reversibility of quantum probabilities

We consider two observers, Alice and Bob, and a single quantum system $\mathcal{S}$. Each of
them can measure a different observable (respectively A and B) of the quantum sys-
tem $\mathcal{S}$ (for simplicity $\mathcal{S}$ is taken to have a finite number of states, i.e. its Hilbert space
will be finite dimensional). We take these observations to be perfect (non demolition)
test measurements, i.e. yes/no measurements, represented by some selfadjoint pro-
jectors $\mathbf{P}_A$ and $\mathbf{P}_B$ such that $\mathbf{P}_A^2 = \mathbf{P}_A$ and $\mathbf{P}_B^2 = \mathbf{P}_B$, but not necessarily commuting. The
eigenvalues of these operators are 1 and 0, corresponding to the two possible out-
comes 1 and 0 (or TRUE and FALSE) of the measurements of the observables A and
of the observable B.

Let us consider now the two following "experimental" protocols, where Alice and
Bob make successive ideal measurements on a system $\mathcal{S}$, and where Alice tries to
guess the result of the measurement by Bob. The two protocols correspond respec-
tively to prediction and to retrodiction.

Protocol 1 - From Alice to Bob: Alice gets the system $\mathcal{S}$ (in a state she knows noth-
ing about). She measures A and if she finds TRUE, then she send the system to Bob, who
measures B. What is the plausibility [4] for Alice that Bob will find that B is TRUE? Let

---

4. In a Bayesian sense.

us call this the conditional probability for B to be found true, A being known to be true, and denote it $P(B \leftarrow\!\!\shortmid A)$. The arrow $\leftarrow\!\!\shortmid$ denotes the causal/time ordering between the measurement of A (by Alice) and of B (by Bob).

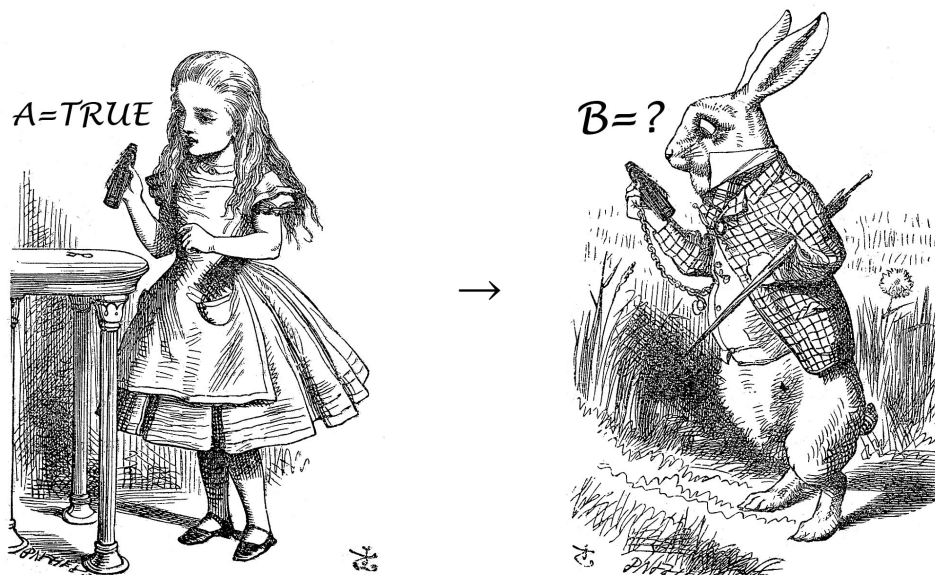

$A = TRUE$ $\rightarrow$ $B = ?$

Figure 2.7: Protocol 1: Alice wants to guess what will be the result of Bob's measurement. This defines the conditional probability $P(B \leftarrow\!\!\shortmid A)$.

Protocol 2 - From Bob to Alice: Alice gets the system $\mathcal{S}$ from Bob, and knows nothing else about $\mathcal{S}$. Bob tells her that he has measured B, but does not tell her the result of his measurement, nor how the system was prepared before he performed the measurement (he may know nothing about it, he just measured B). Then Alice measures A and (if) she finds TRUE she asks herself the following question: what is the plausibility (for her, Alice) that Bob had found that B was TRUE? [5] Let us call this the conditional probability for B to have been found true, A being known to be true, and denote it by $P(B \shortmid\!\!\rightarrow A)$. The arrow $\shortmid\!\!\rightarrow$ denotes the causal/time ordering between the measurement of A (by Alice) and of B (by Bob).

Comparing the two protocols: If $\mathcal{S}$ was a classical system, and the mesurements were classical measurements which do not change the state of $\mathcal{S}$, then the two protocols are equivalent and the two quantities equal the standard conditional probability, given by Bayes formula.

$$\mathcal{S} \text{ classical system}: \quad P(B \leftarrow\!\!\shortmid A) = P(B \shortmid\!\!\rightarrow A) = P(B|A) = P(B \cap A)/P(A).$$

In the quantum case, at a purely logical level, knowing only that the measurement process may perturb the system $\mathcal{S}$, the two conditional probabilities $P(B \leftarrow\!\!\shortmid A)$

---

5. This question makes sense if for instance, Alice has made a bet with Bob. Again, and especially for this protocol, the probability has to be taken in a Bayesian sense.
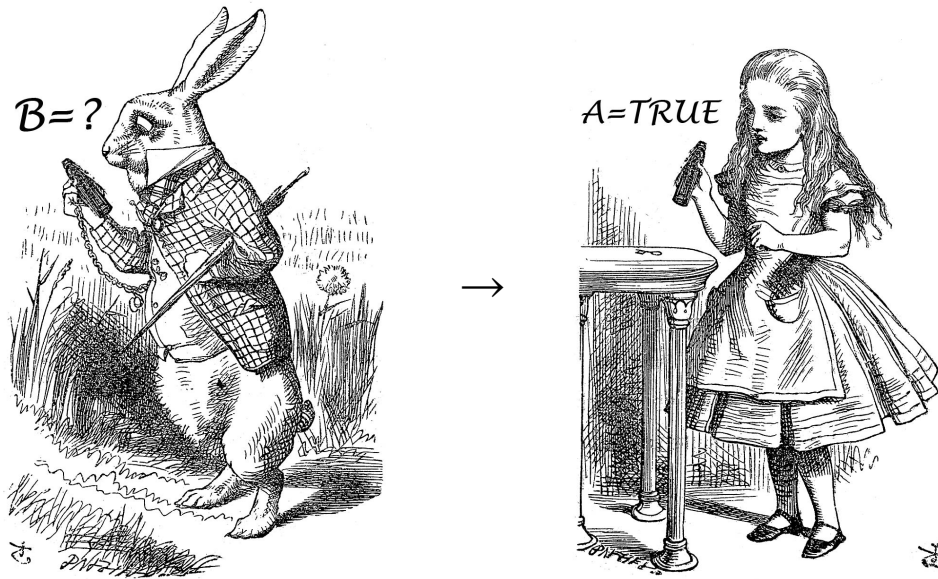
Figure 2.8: Protocol 2: Alice wants to guess what was the result of Bob's measurement. This defines the conditional probability $P(B \mapsto A)$.

and $P(B \mapsto A)$ might very well be different. A crucial and remarkable property of quantum mechanics is that they are still equal.

It is a simple but useful calculation to check this. In the first protocol $P(B \leftarrow A)$ is given by the Born rule; if Alice finds that A is TRUE and knows nothing more, her best bet is that the state of $\mathcal{S}$ is given by the density matrix

$$\rho_A = \mathbf{P}_A / \mathrm{Tr}(\mathbf{P}_A)$$

(equiprobabilities on the eigenspace of P_A with eigenvalue 1, this is already a Bayesian argument, already used by von Neumann [vN32]). Therefore for her the probability for Bob to find that B is TRUE is

$$P(B \leftarrow A) = \mathrm{tr}(\rho_A \mathbf{P}_B).$$

In the second protocol the best guess for Alice is to assume that before Bob measures B the state of the system is given by the fully equidistributed density matrix $\rho_1 = \mathbf{1}/\mathrm{tr}(\mathbf{1})$ (again a Bayesian argument). In this case the probability that Bob finds that B is TRUE, then that Alice finds that A is TRUE, is

$$p_1 = \mathrm{tr}(\mathbf{P}_B)/\mathrm{tr}(\mathbf{1}) \times \mathrm{tr}(\rho_B \mathbf{P}_A) \quad \text{with} \quad \rho_B = \mathbf{P}_B / \mathrm{Tr}(\mathbf{P}_B).$$

Similarly the probability that Bob finds that B is FALSE, then that Alice finds that A is TRUE is

$$p_2 = \mathrm{tr}(\mathbf{1} - \mathbf{P}_B)/\mathrm{tr}(\mathbf{1}) \times \mathrm{tr}(\rho_{\overline{B}} \mathbf{P}_A) = (\mathrm{tr}(\mathbf{P}_A) - \mathrm{tr}(\mathbf{P}_A \mathbf{P}_B))/\mathrm{tr}(\mathbf{1})$$

where $\rho_{\overline{B}} = (\mathbf{1} - \mathbf{P}_B)/\mathrm{tr}(\mathbf{1} - \mathbf{P}_B)$. The total probability is then

$$P(B \mapsto A) = p_1 + p_2 = \mathrm{tr}(\rho_A \mathbf{P}_B).$$

Therefore, even if A and B are not compatible observables, so that the projectors $\mathbf{P}_A$ and $\mathbf{P}_B$ do not commute, one obtains in both case the same, and standard result for quantum conditional probabilities

$$\mathcal{S} \text{ quantum system}: \quad P(B \leftarrow A) = P(B \rightarrow A) = \mathrm{Tr}[\mathbf{P}_A \mathbf{P}_B]/\mathrm{Tr}[\mathbf{P}_A] \qquad (2.5.1)$$

2.5.3 Causal reversibility

This is the basic argument. The situation studied in [ABL64] is more complicated. It involves the selection of some initial state, a series of measurements and the post-selection of some final state, but the conclusion is the same.

This reversibility property of quantum conditional probabilities is very important and is, in my opinion, a crucial feature of quantum mechanics. In this review, I shall denote it *causal reversibility*, in order not to confuse it with time reversal invariance or with the simpler property of reversibility of unitary dynamics.