Lecture 5

QUANTUM MECHANICS

Stéphane ATTAL

Abstract This lecture proposes an introduction to the axioms of Quantum Mechanics and its main ingredients: states, observables, measurement, quantum dynamics. We also discuss how these axioms are changed when considering coupled quantum systems and quantum open systems. This leads in particular to the notion of density matrices, which we explore.

No particular knowledge in Physics is needed in this lecture, as we start from the very beginning in Quantum Mechanics and as our presentation is very axiomatic. Mathematically we make heavy use of the general theory of operators, Spectral Theory, functional calculus for Hermitian operators. These elements can be found in Lecture 1. We also deal a lot with tensor products of Hilbert spaces and with partial traces of operators (see Lecture 2 if necessary).

Stéphane ATTAL
Institut Camille Jordan, Université Lyon 1, France
e-mail: attal@math.univ-lyon1.fr

5.1 The Axioms of Quantum Mechanics

5.1.1 Introduction

The theory of Quantum Mechanics differs a lot from the classical theory of Mechanics that we learned at school. We are used to a theory where systems have a definite position, velocity or energy ... These different characteristics of the system (also called *observables*) can be measured precisely. The result of the measure does not affect the system in general. If the experiment is repeated with exactly the same conditions, it gives exactly the same results.

In the Quantum Mechanics facts are totally different. It is impossible to assign a fixed value for the position, the velocity or the energy of a particle. The state of the particle is a superposition of several possible values (sometimes a continuum of possible values). The system truly occupies all these different values at the same time; for example, when concerning the position, the quantum system can be affected by transformations of the space at different places. In Quantum Mechanics the measurement of some physical quantity concerning a quantum system does not lead to a deterministic value, the result of the measurement is random. Even if the measurement is repeated with exactly the same conditions, the result appears unpredictable. The only quantity which is deterministic and known by the physicist before the measurement process, is the probability distribution of these results. Another fundamental fact is that the effect of measuring the value of a physical parameter of a quantum system (such as position, energy...) affects the system in an irreversible way.

An adequate mathematical language for describing the rules of Quantum Mechanics has been developed about 90 years ago and has shown an incredible efficiency regarding experiments. This axiomatic language, that we develop below, is the theory accepted nowadays by at least 99 % of the physicists. It is not our purpose here to discuss why such incredible axioms represent the reality of the world. Besides, nobody can answer such a question today! We shall take a certain description of the world for granted and we are reassured by 90 years of experiments agreeing with its predictions, to incredible accuracy.

5.1.2 The Axioms

Here are the four main axioms of Quantum Mechanics. They are the only ingredients necessary for being able to compute in Quantum Mechanics (together with a fifth axiom on bipartite quantum system, to be developed later on). This axiomatization does not help much in understanding Quantum Mechanics, nor in making out an intuition on its strange behaviors. One shall only trust the fact that these axioms give rise, since 90 years, to the most accurate results and predictions of all physical theories.

1st Axiom: States

The space of all possible *states* of a quantum system is represented by a Hilbert space $\mathcal{H}$. More precisely the states of a quantum system are *rays* in a Hilbert space. By "rays" we mean equivalence classes of vectors that differ by a non zero scalar multiplication. We can choose a representative vector of that class to have unit norm. These are the so-called *wave functions*, that is, norm 1 elements of $\mathcal{H}$. Note that $\Psi \in \mathcal{H}$ and $e^{i\theta}\Psi$ describe the same state.

A wave function Ψ contains all the possible information about the system; one should be able to compute any parameter of the system from Ψ.

2nd Axiom: Observables

Any physical quantity attached to the quantum system, which can be measured, such as position, velocity, energy, spin, ... is represented by a self-adjoint operator on $\mathcal{H}$. These are called *observables* of the system.

The set of different possible values for the measurement of an observable X is its spectrum $\sigma(\mathsf{X})$. In particular, for some quantum systems, the energy of the system may take values only in a discrete set. This is the origin of the name "Quantum Mechanics": the energy is made of *quanta*, small incompressible values.

Recall that, in the finite dimensional case, every observable X can be diagonalized in some orthonormal basis. This means that X can be written as

$$\mathsf{X} = \sum_{i=1}^{n} \lambda_i \, \mathsf{P}_i$$

where the λ_i are the eigenvalues of X and the P_i are the orthogonal projectors onto the eigenspaces. Note that, for all bounded measurable function f on $\mathbb{R}$ we have

$$f(\mathsf{X}) = \sum_{i=1}^{n} f(\lambda_i)\mathsf{P}_i \, .$$

In the infinite dimensional case, an observable X is described by its spectral measure $\xi_{\mathsf{X}}(\cdot)$ and by the Spectral Theorem:

$$\mathsf{X} = \int_{\sigma(\mathsf{X})} \lambda \; \mathrm{d}\xi_{\mathsf{X}}(\lambda) \, .$$

Recall that for every bounded measurable function f on $\mathbb{R}$ we have

$$f(\mathsf{X}) = \int_{\sigma(\mathsf{X})} f(\lambda) \; \mathrm{d}\xi_{\mathsf{X}}(\lambda) \, .$$

3rd Axiom: Measurement

The only possible numerical outcome for the measurement of an observable X is an element of its spectrum $\sigma(\mathsf{X})$. The result of the measurement belongs to the set $\sigma(\mathsf{X})$ but is completely random. The only possible information one can be sure of is the probability distribution of the measurement outcome. This is described as follows.

Let ξ_{X} be the spectral measure associated to an observable X. If the state of the system is Ψ then the probability of measuring, for the observable X, a value which lies in the Borel set $A \subset \mathbb{R}$ is

$$\boxed{\mathbb{P}(\mathsf{X} \in A) = \|\xi_{\mathsf{X}}(A)\,\Psi\|^2\ ,}$$

which is also equal to

$$\boxed{\mathbb{P}(\mathsf{X} \in A) = \langle \Psi\,,\,\xi_{\mathsf{X}}(A)\,\Psi\rangle}$$

and to

$$\boxed{\mathbb{P}(\mathsf{X} \in A) = \mathrm{Tr}\,\big(|\Psi\rangle\langle\Psi|\,\xi_{\mathsf{X}}(A)\big)\,.}$$

Furthermore, immediately after the measurement, the state of the system is changed and becomes

$$\boxed{\Psi' = \frac{\xi_{\mathsf{X}}(A)\,\Psi}{\|\xi_{\mathsf{X}}(A)\,\Psi\|}\ .}$$

This is the so-called *reduction of the wave packet*.

Note that, in the case where the state space $\mathcal{H}$ is finite dimensional, a measurement of an observable

$$\mathsf{X} = \sum_{i=1}^{n} \lambda_i\,\mathsf{P}_i$$

will give randomly one of its eigenvalues λ_i, as outcome of the measurement, with probability

$$\boxed{\mathbb{P}(\mathsf{X} = \lambda_i) = \langle \Psi\,,\,\mathsf{P}_i\,\Psi\rangle = \|\mathsf{P}_i\,\Psi\|^2\,.}$$

After that measurement the state of the system is changed into

$$\boxed{\Psi' = \frac{\mathsf{P}_i\,\Psi}{\|\mathsf{P}_i\,\Psi\|}\ .}$$

Note the following important fact: when measuring successively with respect to two different observables of a quantum system $\mathcal{H}$ the order in which the measurements are made is of importance in general. For example, assume for simplicity that we are in finite dimension, that the spectral decompositions of the observables X and Y are respectively

$$\mathsf{X} = \sum_{i=1}^{n} \lambda_i \, \mathsf{P}_i \qquad \text{and} \qquad \mathsf{Y} = \sum_{i=1}^{n} \mu_i \, \mathsf{Q}_i \, .$$

Assume that the initial state is Ψ, then the probability of measuring the observable X equal to λ_i and then Y equal to μ_j is given as follows. Measuring first X to be equal to λ_i is obtained with probability $\|\mathsf{P}_i \, \Psi\|^2$ and the state becomes

$$\Psi' = \frac{\mathsf{P}_i \, \Psi}{\|\mathsf{P}_i \, \Psi\|} \, .$$

Having this state now, the measurement of Y being equal to μ_j is obtained with probability

$$\frac{\|\mathsf{Q}_j \, \mathsf{P}_i \, \Psi\|^2}{\|\mathsf{P}_i \, \Psi\|^2} \, ,$$

the final state is then

$$\Psi'' = \frac{\mathsf{Q}_j \, \Psi'}{\|\mathsf{Q}_j \, \Psi'\|} = \frac{\mathsf{Q}_j \, \mathsf{P}_i \, \Psi}{\|\mathsf{Q}_j \, \mathsf{P}_i \, \Psi\|} \, .$$

The total probability of the two successive measurements is then (with obvious notations)

$$\mathbb{P}(\mathsf{X} = \lambda_i \, , \mathsf{Y} = \mu_j) = \mathbb{P}(\mathsf{Y} = \mu_j \, | \, \mathsf{X} = \lambda_i) \, \mathbb{P}(\mathsf{X} = \lambda_i)$$

which gives finally

$$\mathbb{P}(\mathsf{X} = \lambda_i \, , \mathsf{Y} = \mu_j) = \|\mathsf{Q}_j \, \mathsf{P}_i \, \Psi\|^2 \, .$$

Had we done the measurements in the reverse order, we would have obtained the probability $\|\mathsf{P}_i \, \mathsf{Q}_j \, \Psi\|^2$ and the final state

$$\Psi'' = \frac{\mathsf{P}_i \, \mathsf{Q}_j \, \Psi}{\|\mathsf{P}_i \, \mathsf{Q}_j \, \Psi\|} \, .$$

We leave to the reader to check that if the observables X and Y do not commute, that is, if the orthogonal projectors P_i and Q_j do not commute two-by-two, then the two probabilities and the two final states above are different in general.

On the other hand, if X and Y commute, their spectral projectors do commute with each other and hence the two ways of measuring X and Y give the same results, that is, the same probability and the same final state after measurement.

4th Axiom: Dynamics

One observable of the system $\mathcal{H}$ has a particular status: the total energy of the system. This observable, let us denote it by H, is called the *Hamiltonian* of the system, it controls the way the system evolves with time, without exterior intervention. Indeed, if we put for all $t \in \mathbb{R}$

$$\mathsf{U}_t = \mathrm{e}^{-it\mathsf{H}}\,,$$

in the sense of the Functional Calculus for operators, then $(\mathsf{U}_t)_{t\in\mathbb{R}}$ forms a group of unitary operators on $\mathcal{H}$. The state of the system at time t is then given by

$$\boxed{\Psi_t = \mathsf{U}_t\,\Psi_0}$$

if it were Ψ_0 at time 0. This is the so-called *Schrödinger equation.*

Note that the probability of measuring an observable X to be in a set A at time t is thus equal to

$$\|\xi_{\mathsf{X}}(A)\,\Psi_t\|^2 = \langle \Psi_t\,,\,\xi_{\mathsf{X}}(A)\,\Psi_t\rangle = \langle \mathsf{U}_t\,\Psi_0\,,\,\xi_{\mathsf{X}}(A)\,\mathsf{U}_t\,\Psi_0\rangle = \langle \Psi_0\,,\,\mathsf{U}_t^*\,\xi_{\mathsf{X}}(A)\,\mathsf{U}_t\,\Psi_0\rangle\,.$$

But, by the Bounded Functional Calculus, the mapping $A \mapsto \mathsf{U}_t^*\,\xi_{\mathsf{X}}(A)\,\mathsf{U}_t$ is the spectral measure $A \mapsto \xi_{\mathsf{X}_t}(A)$, where X_t is the observable

$$\mathsf{X}_t = \mathsf{U}_t^*\mathsf{X}\mathsf{U}_t\,.$$

Hence this probability is the same as the probability to measure the observable X_t in the state Ψ_0.

This equivalent point of view is the so-called *Heisenberg picture*, as opposed to the first one which is called the *Schrödinger picture.* That is, instead of considering that states are evolving with time and that we always measure the same observable, one can think of the observables evolving with time and the state being fixed.

The Schrödinger picture is that most commonly used by physicists. It is more natural to think the time evolution as modifying the state of the system, the observables being fixed functionals which respect to which we make measures (with a designed apparatus for example).

But when having a probabilistic point of view on this theory, one may find the Heisenberg picture more natural. The state is interpreted as a fixed underlying probability measure, the observables are kind of random variables from which we extract a probability distribution. As time goes the random variables evolve (they are "processes") and their distributions also evolve with time.

5.1.3 Stern-Gerlach Experiment

In order to illustrate the above postulates we describe a well-known experiment, due to Stern and Gerlach.

A source emits a linear horizontal beam of particles (electrons in our example). This beam goes through an intense vertical magnetic field. One observes that the beam splits into several beams corresponding to fixed deviations. The number of such beams is finite, fixed, and depends only on the nature of the particles (2 beams for electrons, for example). We call each of the possible deviations of the particles the *spin* (in the direction Oz). In the case of the electron one speaks for example of spin $+1$ or -1, depending on which of the two deviations occurs. If one makes the particles go through the field one by one we see them "choosing" a spin at random. The value of the spin measured for each particle cannot be predicted. The only predictable fact is that after a large number of particles have gone through the experiment, there will be a fixed and predictable proportion of particles in each direction.

These different deviations actually correspond to a kind of magnetic moment of the particle. The fact that each particle has only a discrete spectrum of spin values is fundamental in quantum mechanics.

Suppose that we isolate a beam of electrons which all have spin $+1$ in the vertical direction. If one makes this beam go through the same vertical field again, one observes that they all deviate corresponding to the spin $+1$ direction. These particles seem to "have kept in mind" that they are spin $+1$.

Now, suppose that this selected beam goes through a magnetic field whose direction makes an angle θ with the vertical axis. One then observes another splitting for the electrons, into two directions with respective proportions $\cos^2(\theta/2)$ and $\sin^2(\theta/2)$ this time.

Finally, we select the spin $+1$ beam after this experiment (with the field in the direction θ) and make it go through the initial vertical field again. One observes a splitting into two directions, with proportions $\cos^2(\theta/2)$ and $\sin^2(\theta/2)$ respectively. Recall that this beam was selected with spin $+1$ in the vertical direction; it appeared to have memory of this fact. We see that, after the passage through the direction θ field, the beam has lost the memory of its spin in the vertical direction.

Trying to give a model for such an experiment and such a complicated behavior may seem very difficult. One might be tempted to attach a random variable to the spin in each direction and to find rules explaining how these random variables are modified after each experiment. This model is clearly, at the least, very complicated, and, in fact, impossible (see Section **??**). However, the formalism of quantum mechanics provides a very simple answer. Let us see how the axioms of quantum mechanics describe in a very simple and nice way the Stern-Gerlach experiment.

In order to give a model for the spin observables it is enough to consider the state space $\mathcal{H} = \mathbb{C}^2$. A state is then an element $\Psi = (u, v) \in \mathbb{C}^2$ with norm 1. The spin observable in the normalized direction (x, y, z) is physically represented by the observable

$$\mathsf{S}_{(x,y,z)} = \begin{pmatrix} z & x - iy \\ x + iy & -z \end{pmatrix}.$$

This operator has eigenvalues $+1$ and -1 with unit eigenvectors α_+ and α_- respectively. In the vertical direction the spin observable is

$$\mathsf{S}_{(0,0,1)} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

with $\alpha_+ = e_1$ and $\alpha_- = e_2$ (the canonical basis vectors). Thus the spin of the particle is $+1$ with probability $p = |\langle \Psi \,,\, \alpha_+ \rangle|^2$ and is -1 with probability $q = |\langle \Psi \,,\, \alpha_- \rangle|^2$. After going through the vertical field, those particles, which have been observed with spin $+1$, are in the state

$$\Psi' = \frac{\langle \Psi \,,\, e_1 \rangle \; e_1}{\|\langle \Psi \,,\, e_1 \rangle \; e_1\|} \,,$$

that is, $\Psi' = e_1$ (recall that wave functions which differ by a modulus one complex factor represent the same state).

If one measures these particle spins in the vertical direction again, we get a probability $p = |\langle e_1 \,,\, \alpha_+ \rangle|^2 = |\langle e_1 \,,\, e_1 \rangle|^2 = 1$ to measure it with spin $+1$; and hence a probability 0 to measure it with spin -1. The beam has "remembered" that it has spin $+1$ in the vertical direction.

If we now measure the spin in the direction $(0, \sin\theta, \cos\theta)$ the spin observable is

$$\mathsf{S}_{(0,\sin\theta,\cos\theta)} = \begin{pmatrix} \cos\theta & -i\sin\theta \\ i\sin\theta & -\cos\theta \end{pmatrix} .$$

The associated eigenvectors are

$$\alpha_+ = \begin{pmatrix} -i\cos(\theta/2) \\ \sin(\theta/2) \end{pmatrix} \qquad \text{and} \qquad \alpha_- = \begin{pmatrix} \sin(\theta/2) \\ i\cos(\theta/2) \end{pmatrix} .$$

In particular, the respective probabilities are

$$p = |\langle e_1 \,,\, \alpha_+ \rangle|^2 = \cos^2(\theta/2) \qquad \text{and} \qquad q = |\langle e_1 \,,\, \alpha_- \rangle|^2 = \sin^2(\theta/2) \,.$$

We recover the observed proportions. Those particles which have spin $+1$ in this direction θ are now in the state $\Psi'' = \alpha_+$.

Finally, if one makes the spin $+1$ particles (measured in the direction θ) go through the vertical field again, we get the proportions $p = |\langle \alpha_+ \,,\, e_1 \rangle|^2 = \cos^2(\theta/2)$ and $q = |\langle \alpha_- \,,\, e_1 \rangle|^2 = \sin^2(\theta/2)$.

This is exactly what was observed. One can only be impressed by the efficiency of this formalism!

5.2 Bipartite Quantum Systems

We need now to extend our axioms to the situation where our quantum system of interest is made of two pieces, two quantum systems. This is a so-called *bipartite quantum system.*

5.2.1 Coupling Quantum Systems

The way one can describe a quantum system made of two quantum subsystems, the way one defines states on the whole system, or the way observables of each subsystem are appearing in the framework of the larger system, is the object of a new axiom. Even though this axiom can somehow be justified if one writes all the mathematical consequences that one may expect from a coupled system: space containing the two initial spaces, independence of the measurement results on each subsystem, etc., we do not wish to develop this justification here and we take this assumption for granted, as a new axiom of quantum mechanics.

5th axiom: Coupled systems

If the respective state space of two quantum systems are $\mathcal{H}$ and $\mathcal{K}$, then the state space of the coupled system is the tensor product

$$\boxed{\mathcal{H} \otimes \mathcal{K}\,.}$$

If the two systems were initially in the respective states ψ and ϕ, then the coupled system is in the state

$$\boxed{\psi \otimes \phi\,.}$$

But be aware that the coupled system $\mathcal{H} \otimes \mathcal{K}$ can be in a more complicated state, that is, any norm one vector $\Psi \in \mathcal{H} \otimes \mathcal{K}$; this means a state where $\mathcal{H}$ and $\mathcal{K}$ are not necessarily independent at the beginning.

If X and Y are observables of the systems $\mathcal{H}$ and $\mathcal{K}$ respectively, then they are extended as observables of $\mathcal{H} \otimes \mathcal{K}$ by considering the operators

$$\boxed{\mathsf{X} \otimes \mathrm{I} \qquad \text{and} \qquad \mathrm{I} \otimes \mathsf{Y}}$$

respectively. They mean the same: if X represents the position of the system $\mathcal{H}$ alone, then $\mathsf{X} \otimes \mathrm{I}$ still represent the position of the particle $\mathcal{H}$, but as part of a larger system now.

It is easy to check that the axioms of Subsection 5.1.2 are preserved when seeing $\mathcal{H}$ as a subsystem of $\mathcal{H} \otimes \mathcal{K}$. Indeed, first note that the spectrum of the operator $\mathsf{X}\otimes\mathrm{I}$ is $\sigma(\mathsf{X})$ and the spectral measure of $\mathsf{X}\otimes\mathrm{I}$ is

$$\xi_{\mathsf{X}\otimes\mathrm{I}}=\xi_{\mathsf{X}}\otimes\mathrm{I}\,,$$

as can be checked easily (exercise). Hence, measuring the observable X of $\mathcal{H}$, when the coupled system $\mathcal{H}\otimes\mathcal{K}$ is in the state $\psi\otimes\phi$, gives the probabilities

$$\begin{aligned}\mathbb{P}(\mathsf{X}\otimes\mathrm{I}\in A)&=\langle\psi\otimes\psi\,,\,\xi_{\mathsf{X}\otimes\mathrm{I}}(A)\,(\psi\otimes\phi)\rangle\\&=\langle\psi\otimes\psi\,,\,(\xi_{\mathsf{X}}(A)\otimes\mathrm{I})\,(\psi\otimes\phi)\rangle\\&=\langle\psi\,,\,\xi_{\mathsf{X}}(A)\,\psi\rangle\,\langle\phi\,,\,\phi\rangle\\&=\langle\psi\,,\,\xi_{\mathsf{X}}(A)\,\psi\rangle\\&=\mathbb{P}(\mathsf{X}\in A)\,.\end{aligned}$$

We recover the usual probabilities for the system $\mathcal{H}$ alone.

The reduction of the wave packet also works smoothly in this enlarged point of view. After having measured the observable $\mathsf{X}\otimes\mathrm{I}$ in the set A, the state of the system becomes

$$\frac{\xi_{\mathsf{X}\otimes\mathrm{I}}(A)\,(\psi\otimes\phi)}{\|\xi_{\mathsf{X}\otimes\mathrm{I}}(A)\,(\psi\otimes\phi)\|}=\frac{\xi_{\mathsf{X}}(A)\psi}{\|\xi_{\mathsf{X}}(A)\psi\|}\otimes\phi\,.$$

We recover the usual reduced state

$$\frac{\xi_{\mathsf{X}}(A)\psi}{\|\xi_{\mathsf{X}}(A)\psi\|}$$

for the system $\mathcal{H}$, tensorized by the unchanged state ϕ of the system $\mathcal{K}$.

This is all for the moment concerning coupled quantum systems. We shall come back to it much more in details when dealing with quantum open systems.

5.2.2 Hidden Variables, Bell's Inequalities

The experiment that we want to describe now is a very famous one which shows that the probabilistic behavior of Quantum Mechanics cannot be modeled by classical Probability Theory.

Consider a system made of two particles of the same nature going in two different directions (right and left, say). Assume that their spin can take only two values: $+1$ and -1. They have been prepared in such a way that their spins, in a fixed direction, are *anticorrelated*. This means the following. The state space of the two particles (just for the study of their respective spin) is $\mathcal{H}=\mathbb{C}^2\otimes\mathbb{C}^2$. If (e_1,e_2) is an orthonormal basis of $\mathbb{C}^2$ made of spin $+1$ and spin -1 eigenvectors respectively (in the chosen direction), then the initial state of our system is

$$\varphi = \frac{1}{\sqrt{2}}(e_1 \otimes e_2 - e_2 \otimes e_1).$$

Actually, the above state does not depend on the choice of direction: as a state (that is, up to a phase factor) it is the same if one changes the choice of the orthonormal basis. The above state is called the *singlet state* in physics: it is the state in which the spins of the two particles are anticorrelated in any direction of the space. Such a state is physically realizable.

In front of each particle (right and left) is placed an apparatus which measures their spin in directions $(\sin\alpha, 0, \cos\alpha)$ and $(\sin\beta, 0, \cos\beta)$ respectively (that we shall call directions α and β in the following). Let us denote by X_α the observable "spin of the left particle in the direction α", and by Y_β the observable "spin of the right particle in the direction β". They are observables of the first and the second system respectively. That is, if one wants to consider them as observables of the whole system, we have to consider the observables $\mathsf{X}_\alpha \otimes \mathrm{I}$ and $\mathrm{I} \otimes \mathsf{Y}_\beta$ respectively.

These two observables obviously commute hence, when measuring both of them, the order of measurement is of no importance (as was discussed after the 3rd Axiom).

We denote by

$$\begin{aligned}
\alpha_+ &= (\cos\alpha/2, \sin\alpha/2)\\
\alpha_- &= (-\sin\alpha/2, \cos\alpha/2)\\
\beta_+ &= (\cos\beta/2, \sin\beta/2)\\
\beta_- &= (-\sin\beta/2, \cos\beta/2)
\end{aligned}$$

the eigenvectors of X_α and Y_β respectively. The probability that the spin is measured to be $+1$ on the left and $+1$ on the right is equal to

$$\begin{aligned}
\mathbb{P}(\mathsf{X}_\alpha = +1, \mathsf{Y}_\beta = +1) &= \left\| (|\alpha_+\rangle\langle\alpha_+| \otimes \mathrm{I})\,(\mathrm{I} \otimes |\beta_+\rangle\langle\beta_+|)\,\phi \right\|^2\\
&= \left\| |\alpha_+ \otimes \beta_+\rangle\langle\alpha_+ \otimes \beta_+|\,\varphi \right\|^2\\
&= \left|\langle\alpha_+ \otimes \beta_+\,,\,\varphi\rangle\right|^2 .
\end{aligned}$$

A straightforward computation then gives

$$\mathbb{P}(\mathsf{X}_\alpha = +1, \mathsf{Y}_\beta = +1) = \frac{1}{2}\sin^2\left(\frac{\alpha - \beta}{2}\right) .$$

In the same way we find

$$\mathbb{P}(\mathsf{X}_\alpha = +1, \mathsf{Y}_\beta = -1) = \frac{1}{2}\cos^2\left(\frac{\alpha-\beta}{2}\right)$$
$$\mathbb{P}(\mathsf{X}_\alpha = -1, \mathsf{Y}_\beta = +1) = \frac{1}{2}\cos^2\left(\frac{\alpha-\beta}{2}\right)$$
$$\mathbb{P}(\mathsf{X}_\alpha = -1, \mathsf{Y}_\beta = -1) = \frac{1}{2}\sin^2\left(\frac{\alpha-\beta}{2}\right).$$

We wonder if such correlations, such probabilities, can be obtained with the help of a classical probability model. That is, we wonder if it is possible to define a probability space $(\Omega, \mathcal{F}, P)$ and some ± 1-valued random variables X_α, Y_β on $(\Omega, \mathcal{F}, P)$, for each angle α and β, such that the above is satisfied. This assumption would mean physically that there is some uncertainty in the knowledge of the initial state of the system; everything is determined from the beginning (the spin in each direction, etc...) but we have a lack of knowledge in some of the variables of the system which leads to a random result about the quantities X_α, Y_β (this is the so-called *hidden variable* hypothesis).

We then conduct the following experiment. We are given three fixed angles α_1, α_2 and α_3. We have a great number of particle pairs all prepared in the same state as described above. For each choice of a pair (α_i, α_j) we make a large number of our particle pairs going through a spin measurement on the left and on the right particle. We then get all the different correlations as computed above.

But in fact we have the following easy result.

Theorem 5.1. [Bell's Three Variable Inequality] *For any three ± 1-valued random variables X_1, X_2, X_3 on a probability space $(\Omega, \mathcal{F}, P)$ we have*

$$\mathbb{P}(X_1 = 1, X_3 = -1) \leq \mathbb{P}(X_1 = 1, X_2 = -1) + \mathbb{P}(X_2 = 1, X_3 = -1).$$

Proof. Simply write

$$\begin{aligned}\mathbb{P}(X_1 = 1, &X_3 = -1) = \\ &= \mathbb{P}(X_1 = 1, X_2 = -1, X_3 = -1) + \mathbb{P}(X_1 = 1, X_2 = 1, X_3 = -1) \\ &\leq \mathbb{P}(X_1 = 1, X_2 = -1) + \mathbb{P}(X_2 = 1, X_3 = -1).\end{aligned}$$

□

With the probabilities we found, Bell's inequality is violated. For example take $\alpha_1 = \pi/2$, $\alpha_2 = 5\pi/6$ and $\alpha_3 = 7\pi/6$, we find

$$\mathbb{P}(X_1 = 1, X_3 = -1) = \frac{3}{8}$$

and

$$\mathbb{P}(X_1 = 1, X_2 = -1) + \mathbb{P}(X_2 = 1, X_3 = -1) = \frac{1}{8} + \frac{1}{8}.$$

It is impossible to attach classical random variables behind each spin of the particle. The theory and the experiment (which was performed in Orsay by A. Aspect's team) show that these correlations cannot come from classical random variables.

A possible criticism of the above conclusion is to say that the measurement on the left device influences the measurement on the right one. The Orsay experiment, directed by A. Aspect in 1982, was actually a little more sophisticated than only checking the correlations above. A random choice of two different angles α_1, α_2 was made on the left, and the same with β_1, β_2 on the right.

We then use an improved form of the Bell Inequalities.

Theorem 5.2 (Bell's Four Variable Inequality). *For any quadruple X_1, X_2, Y_1, Y_2 of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, taking the only values -1 and 1, we have*

$$P(X_1 = Y_1) \leq P(X_1 = Y_2) + P(X_2 = Y_1) + P(X_2 = Y_2)\,.$$

Proof. Note that

$$\begin{aligned}\|X_1 + Y_1\|_{L^1(\Omega)} &= \mathbb{E}\left[|X_1 + Y_1|\right] = \mathbb{E}\left[|X_1 + Y_1|\, \mathbb{1}_{X_1 = Y_1}\right] + \mathbb{E}\left[|X_1 + Y_1|\, \mathbb{1}_{X_1 \neq Y_1}\right] \\ &= 2\,\mathbb{E}\left[\mathbb{1}_{X_1 = Y_1}\right] \\ &= 2\,\mathbb{P}(X_1 = Y_1)\,.\end{aligned}$$

Thus the inequality of the theorem is just the quadrangle inequality:

$$\|X_1 + Y_1\|_{L^1(\Omega)} \leq \|X_1 + Y_2\|_{L^1(\Omega)} + \|X_2 + Y_2\|_{L^1(\Omega)} + \|X_2 + Y_1\|_{L^1(\Omega)}\,. \quad \square$$

In our case, we have

$$\mathbb{P}(X_{\alpha_i} = Y_{\beta_j}) = \sin^2\left(\frac{\alpha_i - \beta_j}{2}\right).$$

Thus the Bell inequality is violated for the choice $\alpha_1 = 0$, $\alpha_2 = 2\pi/3$, $\beta_1 = \pi$ and $\beta_2 = \pi/3$, for we find

$$\mathbb{P}(X_1 = Y_1) = 1$$

and

$$\mathbb{P}(X_1 = Y_2) + \mathbb{P}(X_2 = Y_1) + \mathbb{P}(X_2 = Y_2) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4}\,.$$

These results go against the hidden variable hypothesis and the attempt to model Quantum Mechanics with classical Probability Theory. A classical probabilistic model of the phenomena of Quantum Mechanics is not possible.

One comment must be added here. The above argument works if one can be sure that the measurement of the left particle in the direction α cannot influence the measurement of the right particle, a little later, in the direction

β. This assumption is called *locality* in physics. In the Orsay experiment, this assumption was fulfilled. Indeed, they performed the measurements on the left and then on the right in an interval of time which is smaller than the flight time of a photon between the two pieces of apparatus. Thus, by the causality principle of Relativity Theory, the first measurement cannot have influenced the second one.

If one admits the causality principle to be valid in Quantum Mechanics, then the hidden variable hypothesis must be rejected, as well as any attempt to model Quantum Mechanics with classical Probability Theory.

If the causality principle in Quantum Mechanics is abandoned then there is still some room for modeling the Orsay experiment in classical probabilist terms. This is, for example, the case of so-called Bohmian quantum mechanics. But a very large majority of physicists nowadays considers the causality principle to be valid in Quantum Mechanics.

5.3 Quantum Open Systems

The formalism of Quantum Mechanics which is developed in Section 5.1 actually needs to be extended. Indeed, this formalism only describes the behavior of isolated quantum systems, but in many situations one has to consider quantum systems which interact with another one (or other ones). These are the so-called *open quantum systems*.

The main point with open quantum systems, which differs from just considering bipartite quantum systems such as in Subsection 5.2.2, is the fact that in most of the situations of open quantum systems one does not have access to the second system. In general this happens because this second system is too large, too complicated, like a large environment, a quantum heat bath etc.

Even if one is not willing to, it is very difficult to prevent a quantum system to interact with an exterior system, with the environment, with photons ... This fact induces the so-called *decoherence* on the small system, this effect is the main obstacle which prevents physicist from being able to built a quantum computer (at the time we write this book). It is one of the main stream of research in Quantum Physics nowadays to try to prevent, or to control, this decoherence, this interaction with the environment.

It happens also in many situations, in particular in the context of Quantum Information Theory, that one is in the following situation: a quantum system is made of two simple systems which are interacting (or which have been interacting), but the two systems ($\mathcal{A}$ and $\mathcal{B}$ say) are shared by two distant persons. The first person (always called Alice!) has only access to system $\mathcal{A}$, while system $\mathcal{B}$ is the property of Bob. This is also a very important and common situation of open quantum system.

Describing the rules of Quantum Mechanics in the situation of open systems leads to important extensions of the axioms.

5.3.1 Density Matrices

First of all, let us focus on states. We have already seen that a coupled system in Quantum Mechanics is represented by the tensor product of the corresponding Hilbert spaces $\mathcal{H} \otimes \mathcal{K}$. A state on that system is thus a unit vector $\Psi \in \mathcal{H} \otimes \mathcal{K}$.

Imagine that we are dealing with the coupled system $\mathcal{H} \otimes \mathcal{K}$, but we personally have access to $\mathcal{H}$ only. If we want to measure an observable X of $\mathcal{H}$, we have to consider the observable $\mathsf{X} \otimes \mathrm{I}$ on $\mathcal{H} \otimes \mathcal{K}$, as we have seen in the 5th Axiom. The spectral measure of $\mathsf{X} \otimes \mathrm{I}$ is $\xi_{\mathsf{X}}(\cdot) \otimes \mathrm{I}$ and the probability for finding the value of the measurement lying in the Borel set A is equal to

$$\begin{aligned}\mathbb{P}(\mathsf{X} \otimes \mathrm{I} \in A) &= \langle \Psi \,,\, (\xi_{\mathsf{X}}(A) \otimes \mathrm{I})\, \Psi \rangle \\ &= \mathrm{Tr}\, \big(|\Psi\rangle\langle\Psi|\, (\xi_{\mathsf{X}}(A) \otimes \mathrm{I}) \big) \\ &= \mathrm{Tr}\, \big(\mathrm{Tr}_{\mathcal{K}}\, (|\Psi\rangle\langle\Psi|)\ \xi_{\mathsf{X}}(A) \big)\end{aligned}$$

where $\mathrm{Tr}_{\mathcal{K}}$ denotes the partial trace with respect to the space $\mathcal{K}$, and where we have used the basic property characterizing the partial traces. This means that

$$\boxed{\mathbb{P}(\mathsf{X} \otimes \mathrm{I} \in A) = \mathrm{Tr}\, (\rho_{\mathcal{H}}\, \xi_{\mathsf{X}}(A))} \tag{5.1}$$

where

$$\boxed{\rho_{\mathcal{H}} = \mathrm{Tr}_{\mathcal{K}}\, (|\Psi\rangle\langle\Psi|)\ .}$$

Clearly, any physical measurement on any observable of $\mathcal{H}$ will be obtained by the same kind of formula as (5.1).

This formula is very interesting because, despite the fact that one does not have access to the whole system $\mathcal{H} \otimes \mathcal{K}$, despite the fact that one may not know the full state Ψ, if we have the knowledge of this operator $\rho_{\mathcal{H}}$ we can compute any probability on any observable of $\mathcal{H}$, via Formula (5.1). The point with this formula is that it makes use only of operators on $\mathcal{H}$. It takes into account the fact that the system $\mathcal{H}$ is not isolated, but the formula stays internal to $\mathcal{H}$.

Note that one recovers the special case of an isolated system (Subsection 5.1.2) when $\rho_{\mathcal{H}}$ is of the form $|\Phi\rangle\langle\Phi|$.

We now wish to characterize those operators $\rho_{\mathcal{H}}$ on $\mathcal{H}$ that can be obtained this way.

Theorem 5.3. *Let ρ be an operator on a Hilbert space $\mathcal{H}$. Then the following assertions are equivalent.*

1) There exists a Hilbert space $\mathcal{K}$ and a unit vector Ψ in $\mathcal{H}\otimes\mathcal{K}$ such that $\rho = \mathrm{Tr}_{\mathcal{K}}(|\Psi\rangle\langle\Psi|)$ *.*
2) The operator ρ is positive, trace-class and $\mathrm{Tr}\ \rho = 1$ *.*

Proof. If ρ is of the form described in 1) then, for every $\phi\in\mathcal{H}$, we have

$$\begin{aligned}\langle \phi\,,\,\rho\,\phi\rangle &= \mathrm{Tr}\,\big(|\phi\rangle\langle\phi|\ \mathrm{Tr}_{\mathcal{K}}(|\Psi\rangle\langle\Psi|)\big)\\ &= \mathrm{Tr}\,\big((|\phi\rangle\langle\phi|\otimes \mathrm{I})\,|\Psi\rangle\langle\Psi|\big)\\ &= \langle \Psi\,,\,(|\phi\rangle\langle\phi|\otimes\mathrm{I})\,\Psi\rangle\ .\end{aligned}$$

Decomposing $\Psi=\sum_{i,j}\Psi_{ij}\,e_i\otimes f_j$ in some orthonormal basis of $\mathcal{H}\otimes\mathcal{K}$, gives

$$\begin{aligned}\langle \phi\,,\,\rho\,\phi\rangle &= \sum_{ij,k}\overline{\Psi_{ij}}\,\Psi_{kj}\ \langle e_i\,,\,\phi\rangle\ \langle\phi\,,\,e_k\rangle\\ &= \sum_j\left|\sum_i\overline{\Psi_{ij}}\ \langle e_i\,,\,\phi\rangle\right|^2\ \geq 0\,.\end{aligned}$$

This shows that ρ is a positive operator. Furthermore, for any orthonormal basis (e_n) of $\mathcal{H}$ we have

$$\begin{aligned}\mathrm{Tr}\,(\rho) &= \sum_n\langle e_n\,,\,\rho\,e_n\rangle\\ &= \sum_n \mathrm{Tr}\,(|e_n\rangle\langle e_n|\ \mathrm{Tr}_{\mathcal{K}}\,(|\Psi\rangle\langle\Psi|))\\ &= \sum_n \mathrm{Tr}\,((|e_n\rangle\langle e_n|\otimes\mathrm{I})\ |\Psi\rangle\langle\Psi|)\\ &= \sum_n\langle\Psi\,,\,(|e_n\rangle\langle e_n|\otimes\mathrm{I})\ \Psi\rangle\\ &= \left\langle\Psi\,,\,\left(\sum_n|e_n\rangle\langle e_n|\otimes\mathrm{I}\right)\,\Psi\right\rangle\\ &= \langle\Psi\,,\,\Psi\rangle\\ &= 1\,.\end{aligned}$$

This proves that ρ is trace-class and that $\mathrm{Tr}\ \rho = 1$. We have proved the theorem in one direction.

Conversely, if ρ is of the form 2) then, by the fundamental decomposition of trace-class operators, it can decomposed as $\rho=\sum_i\lambda_i\,|e_i\rangle\langle e_i|$, for some orthonormal basis (e_i) of $\mathcal{H}$ and some positive scalars λ_i satisfying $\sum_i\lambda_i=1$.

Put $\mathcal{K}=\mathcal{H}$, the vector $\Psi=\sum_i\sqrt{\lambda_i}\ e_i\otimes e_i$ is norm 1 in $\mathcal{H}\otimes\mathcal{K}$. Furthermore the partial trace $\mathrm{Tr}_{\mathcal{K}}(|\Psi\rangle\langle\Psi|)$ can be easily computed:

$$\mathrm{Tr}_{\mathcal{K}}(|\Psi\rangle\langle\Psi|)=\sum_i\lambda_i\,|e_i\rangle\langle e_i| = \rho\,.$$

We have proved the theorem. □

5.3.2 Extended Axioms

The results obtained above mean that considering open quantum systems leads to a generalization of the notion of state.

1st Extended Axiom: States and Measurement of Observables

A quantum *state* on a general Hilbert space $\mathcal{H}$ is a trace-class, positive operator ρ such that $\mathrm{Tr}\,\rho = 1$. The probability of measuring a numerical outcome for an observable X in the set A is given by the formula

$$\boxed{\mathbb{P}(\mathsf{X} \in A) = \mathrm{Tr}\,\left(\rho\,\xi_{\mathsf{X}}(A)\right)\,.} \tag{5.2}$$

Such operators ρ are called *density matrices*. They can always be decomposed as

$$\boxed{\rho = \sum_{n\in\mathbb{N}} \lambda_n\,|e_n\rangle\langle e_n|} \tag{5.3}$$

for some orthonormal basis (e_n) of eigenvectors and some positive eigenvalues λ_n satisfying $\sum_{n\in\mathbb{N}} \lambda_n = 1$. The set of quantum states (i.e. density matrices) on $\mathcal{H}$ is denoted by $\mathcal{S}(\mathcal{H})$.

The special case where

$$\rho = |\Psi\rangle\langle\Psi|$$

for some norm 1 vector Ψ of $\mathcal{H}$ corresponds to the case where $\mathcal{H}$ is an isolated quantum system. In that case ρ is called a *pure state*. We often identify the rank one projector $|\Psi\rangle\langle\Psi|$ with the wave function Ψ.

Recall that quantum states were defined up to a norm 1 multiplicative constant λ. This now appears more clear if one thinks in terms of the pure state $|\Psi\rangle\langle\Psi|$ instead of the vector Ψ. Indeed the pure state associated to $\lambda\Psi$ is

$$|\lambda\Psi\rangle\langle\lambda\Psi| = \overline{\lambda}\,\lambda\,|\Psi\rangle\langle\Psi| = |\Psi\rangle\langle\Psi|\,.$$

Now let us see how the reduction of the wave packet postulate is modified when considering open quantum systems and density matrices.

Proposition 5.4. *Consider a quantum system $\mathcal{H}$ in a state ρ. Let X be an observable of $\mathcal{H}$. When measuring the observable X we get the measurement value to belong to a set A with probability*

$$\mathbb{P}(\mathsf{X} \in A) = \mathrm{Tr}\,(\rho\,\xi_{\mathsf{X}}(A))\,.$$

Immediately after the measurement the state of the system becomes

$$\rho' = \frac{\xi_{\mathsf{X}}(A)\,\rho\,\xi_{\mathsf{X}}(A)}{\mathrm{Tr}\,(\rho\,\xi_{\mathsf{X}}(A))}\,. \tag{5.4}$$

Proof. By Theorem 5.3 there exists a Hilbert space $\mathcal{K}$ and a pure state Ψ on $\mathcal{H}\otimes\mathcal{K}$ such that $\rho = \mathrm{Tr}_{\mathcal{K}}(|\Psi\rangle\langle\Psi|)$. We have already proved above that the observable $\mathsf{X}\otimes\mathrm{I}$ is then measured to belong to A with probability $\mathrm{Tr}\,(\rho\,\xi_{\mathsf{X}}(A))$.

The reduction of the wave packet postulate says that after that measurement the state of the system $\mathcal{H}\otimes\mathcal{K}$ is

$$\Psi' = \frac{(\xi_{\mathsf{X}}(A)\otimes\mathrm{I})|\Psi\rangle}{\|(\xi_{\mathsf{X}}(A)\otimes\mathrm{I})|\Psi\rangle\|}\,.$$

On the space $\mathcal{H}$ the resulting density matrix is then

$$\rho' = \mathrm{Tr}_{\mathcal{K}}\,(|\Psi'\rangle\langle\Psi'|)\,.$$

Applying the different properties of partial traces we get

$$\begin{aligned}
\rho' &= \frac{1}{\|(\xi_{\mathsf{X}}(A)\otimes\mathrm{I})\,|\Psi\rangle\|^2}\;\mathrm{Tr}_{\mathcal{K}}\big((\xi_{\mathsf{X}}(A)\otimes\mathrm{I})\,|\Psi\rangle\langle\Psi|\,(\xi_{\mathsf{X}}(A)\otimes\mathrm{I})\big)\\
&= \frac{1}{\mathrm{Tr}\,\big((\xi_{\mathsf{X}}(A)\otimes\mathrm{I})\,|\Psi\rangle\langle\Psi|\,(\xi_{\mathsf{X}}(A)\otimes\mathrm{I})\big)}\,\xi_{\mathsf{X}}(A)\;\mathrm{Tr}_{\mathcal{K}}\,(|\Psi\rangle\langle\Psi|)\;\xi_{\mathsf{X}}(A)\\
&= \frac{1}{\mathrm{Tr}\,\big(\,\mathrm{Tr}_{\mathcal{K}}\,\big((\xi_{\mathsf{X}}(A)\otimes\mathrm{I})\,|\Psi\rangle\langle\Psi|\,(\xi_{\mathsf{X}}(A)\otimes I)\big)\big)}\,\xi_{\mathsf{X}}(A)\,\rho\,\xi_{\mathsf{X}}(A)\\
&= \frac{1}{\mathrm{Tr}\,(\xi_{\mathsf{X}}(A)\,\rho\,\xi_{\mathsf{X}}(A))}\,\xi_{\mathsf{X}}(A)\,\rho\,\xi_{\mathsf{X}}(A)\,.
\end{aligned}$$

This gives (5.4). □

This leads to the following extended axiom for the wave packet reduction.

2nd Extended Axiom: Wave Packet Reduction

If a quantum system $\mathcal{H}$ is in the state ρ then the measurement of any observable X of $\mathcal{H}$ in the set A gives rise to a new state for $\mathcal{H}$:

$$\boxed{\rho' = \frac{\xi_{\mathsf{X}}(A)\,\rho\,\xi_{\mathsf{X}}(A)}{\mathrm{Tr}\,(\rho\,\xi_{\mathsf{X}}(A))}\,.} \tag{5.5}$$

In particular, if $\mathcal{H}$ is finite dimensional and if the spectral decomposition of X is

$$\mathsf{X} = \sum_{i=1}^{n} \lambda_i\,\mathsf{P}_i$$

then a measurement of X being equal to λ_i gives rise to the new state

$$\boxed{\rho' = \frac{\mathsf{P}_i\,\rho\,\mathsf{P}_i}{\mathrm{Tr}\,(\rho\,\mathsf{P}_i)}\,.} \tag{5.6}$$

Note that if ρ is the pure state associated to Ψ, then ρ' is the pure state associated to $\mathsf{P}_i\,\Psi/\,\|\mathsf{P}_i\,\Psi\|$, as in the first series of axioms.

Let us see how the last axiom is modified when considering density matrices instead of wave functions. Actually we are only going to develop a very simple situation here: the case where the two coupled systems $\mathcal{H}$ and $\mathcal{K}$ evolve independently. More general evolutions for the coupled system $\mathcal{H}\otimes\mathcal{K}$ are considered in Lecture **??**.

On the coupled system $\mathcal{H}\otimes\mathcal{K}$, we consider an Hamiltonian of the form

$$\mathsf{H} = \mathsf{H}_{\mathcal{H}} \otimes \mathrm{I} + \mathrm{I} \otimes \mathsf{H}_{\mathcal{K}}\,,$$

resulting from the parallel evolution of the two systems, each one having its own Hamiltonian (they do not interact with each other). It is easy to check that the associated unitary group is then $\mathsf{U}(t) = \mathsf{U}_{\mathcal{H}}(t)\otimes\mathsf{U}_{\mathcal{K}}(t)$ where $\mathsf{U}_{\mathcal{H}}(t) = \exp(-\mathrm{i}t\mathsf{H}_{\mathcal{H}})$ and $\mathsf{U}_{\mathcal{K}}(t) = \exp(-\mathrm{i}t\mathsf{H}_{\mathcal{K}})$.

Consider an initial state Ψ on $\mathcal{H}\otimes\mathcal{K}$, it gives rise to a density matrix $\rho_{\mathcal{H}} = \mathrm{Tr}_{\mathcal{K}}\,(\,|\Psi\rangle\langle\Psi|\,)$ on $\mathcal{H}$. The state Ψ evolves into the state $\Psi(t) = \mathsf{U}(t)\Psi$ at time t.

Proposition 5.5. *The density matrix corresponding to the state on $\mathcal{H}$ at time t is given by*

$$\rho_{\mathcal{H}}(t) = \mathsf{U}_{\mathcal{H}}(t)\,\rho_{\mathcal{H}}\,\mathsf{U}_{\mathcal{H}}(t)^*\,. \tag{5.7}$$

Proof. The state Ψ evolves to $\Psi(t) = \mathsf{U}(t)\,\Psi$ at time t. Hence the corresponding density matrix on $\mathcal{H}$ is $\rho_{\mathcal{H}}(t) = \mathrm{Tr}_{\mathcal{K}}(|\Psi(t)\rangle\langle\Psi(t)|)$. In particular, for all $\mathsf{X}\in\mathcal{B}(\mathcal{H})$ we have

$$\begin{aligned}
\mathrm{Tr}\,\left(\rho_{\mathcal{H}}(t)\,\mathsf{X}\right) &= \mathrm{Tr}\,\left(\mathrm{Tr}_{\mathcal{K}}\,(|\Psi(t)\rangle\langle\Psi(t)|)\,\mathsf{X}\right)\\
&= \mathrm{Tr}\,\left(\mathrm{Tr}_{\mathcal{K}}\,(\mathsf{U}(t)\,|\Psi\rangle\langle\Psi|\,\mathsf{U}(t)^*)\,\mathsf{X}\right)\\
&= \mathrm{Tr}\,\left(\mathsf{U}(t)\,|\Psi\rangle\langle\Psi|\,\mathsf{U}(t)^*\,(\mathsf{X}\otimes\mathrm{I})\right)\\
&= \mathrm{Tr}\,\left(|\Psi\rangle\langle\Psi|\,(\mathsf{U}_{\mathcal{H}}(t)^*\otimes\mathsf{U}_{\mathcal{K}}(t)^*)\,(\mathsf{X}\otimes\mathrm{I})\,(\mathsf{U}_{\mathcal{H}}(t)\otimes\mathsf{U}_{\mathcal{K}}(t))\right)\\
&= \mathrm{Tr}\,\left(|\Psi\rangle\langle\Psi|\,(\mathsf{U}_{\mathcal{H}}(t)^*\,\mathsf{X}\,\mathsf{U}_{\mathcal{H}}(t))\otimes\mathrm{I}\right)\\
&= \mathrm{Tr}\,\left(\mathrm{Tr}_{\mathcal{K}}\,(|\Psi\rangle\langle\Psi|)\;(\mathsf{U}_{\mathcal{H}}(t)^*\,\mathsf{X}\,\mathsf{U}_{\mathcal{H}}(t))\right)\\
&= \mathrm{Tr}\,\left(\mathsf{U}_{\mathcal{H}}\;\mathrm{Tr}_{\mathcal{K}}\,(|\Psi\rangle\langle\Psi|)\;\mathsf{U}_{\mathcal{H}}(t)^*\,\mathsf{X}\right)\,.
\end{aligned}$$

As this is true for all $\mathsf{X}\in\mathcal{B}(\mathcal{H})$ this proves that

$$\rho_{\mathcal{H}}(t) = \mathsf{U}_{\mathcal{H}}(t)\,\rho_{\mathcal{H}}\,\mathsf{U}_{\mathcal{H}}(t)^*\,. \qquad \square$$

This is how the evolution of states as described in Subsection 5.1.2 should be generalized when dealing with density matrices.

3rd Extended Axiom: Time Evolution

If a quantum system $\mathcal{H}$ has an initial state described by the density matrix ρ and if it evolves independently of the environment with the Hamiltonian H, then the state of $\mathcal{H}$ at time t is the density matrix

$$\boxed{\rho(t) = \mathsf{U}_t\, \rho\, \mathsf{U}_t^*}$$

where $\mathsf{U}_t = \mathrm{e}^{-\mathrm{i}t\mathsf{H}}$.

Note that if ρ is the pure state associated to Ψ, then ρ' is the pure state associated to $\mathsf{U}_t\,\Psi$, as in the first series of axioms.

5.3.3 Ambiguity of the Purification

Coming back to the characterization of density matrices: they all are partial trace of some pure state on some larger space $\mathcal{H}\otimes\mathcal{K}$. Finding a pure state Ψ on $\mathcal{H}\otimes\mathcal{K}$ such that

$$\rho = \mathrm{Tr}_{\mathcal{K}}\,(|\psi\rangle\langle\psi|)$$

is called a *purification* of ρ. It is easy to see that such a purification is not unique, there are plenty of possibilities. The following theorem (important in Quantum Information Theory) classifies all the possible purifications of ρ.

Theorem 5.6 (GHJW Theorem). *Two unit vectors Ψ and Ψ' of $\mathcal{H}\otimes\mathcal{K}$ satisfy*

$$\mathrm{Tr}_{\mathcal{K}}\,(|\Psi\rangle\langle\Psi|) = \mathrm{Tr}_{\mathcal{K}}\,(|\Psi'\rangle\langle\Psi'|)\ ,$$

if and only if there exists a unitary operator U on $\mathcal{K}$ such that

$$\Psi' = (\mathrm{I}\otimes\mathsf{U})\,\Psi\,.$$

Proof. One direction is easy, if $\Psi' = (\mathrm{I}\otimes\mathsf{U})\,\Psi$ then

$$\mathrm{Tr}_{\mathcal{K}}\,(|\Psi'\rangle\langle\Psi'|) = \mathrm{Tr}_{\mathcal{K}}\,((\mathrm{I}\otimes\mathsf{U})\,|\Psi\rangle\langle\Psi|\,(\mathrm{I}\otimes\mathsf{U}^*))\ .$$

By definition of the partial trace, this means that for all $\mathsf{X}\in\mathcal{B}(\mathcal{H})$ we have

$$\begin{aligned}
\mathrm{Tr}\,(\mathrm{Tr}_{\mathcal{K}}\,(|\Psi'\rangle\langle\Psi'|)\ \mathsf{X}) &= \mathrm{Tr}\,(\mathrm{Tr}_{\mathcal{K}}\,((\mathrm{I}\otimes\mathsf{U})\,|\Psi\rangle\langle\Psi|\,(\mathrm{I}\otimes\mathsf{U}^*))\ \mathsf{X})\\
&= \mathrm{Tr}\,((\mathrm{I}\otimes\mathsf{U})\,|\Psi\rangle\langle\Psi|\,(\mathrm{I}\otimes\mathsf{U}^*)\,(\mathsf{X}\otimes\mathrm{I}))\\
&= \mathrm{Tr}\,(|\Psi\rangle\langle\Psi|\,(\mathrm{I}\otimes\mathsf{U}^*)\,(\mathsf{X}\otimes\mathrm{I})\,(\mathrm{I}\otimes\mathsf{U}))\\
&= \mathrm{Tr}\,(|\Psi\rangle\langle\Psi|\,(\mathrm{I}\otimes\mathsf{U}^*)\,(\mathrm{I}\otimes\mathsf{U})\,(\mathsf{X}\otimes\mathrm{I}))\\
&= \mathrm{Tr}\,(|\Psi\rangle\langle\Psi|\,(\mathsf{X}\otimes\mathrm{I}))\\
&= \mathrm{Tr}\,(\mathrm{Tr}_{\mathcal{K}}\,(|\Psi\rangle\langle\Psi|)\ \mathsf{X})\ .
\end{aligned}$$

As this holds for all $\mathsf{X}\in\mathcal{B}(\mathcal{H})$, we have proved that

$$\mathrm{Tr}_{\mathcal{K}}\left(|\Psi\rangle\langle\Psi|\right) = \mathrm{Tr}_{\mathcal{K}}\left(|\Psi'\rangle\langle\Psi'|\right) .$$

We now prove the converse direction, assume that

$$\mathrm{Tr}_{\mathcal{K}}\left(|\Psi\rangle\langle\Psi|\right) = \mathrm{Tr}_{\mathcal{K}}\left(|\Psi'\rangle\langle\Psi'|\right) .$$

Let $\rho = \mathrm{Tr}_{\mathcal{K}}\left(|\Psi\rangle\langle\Psi|\right)$, it is a density matrix on $\mathcal{H}$. Hence it can be diagonalized in some orthonormal basis of $\mathcal{H}$:

$$\rho = \sum_{n\in\mathcal{N}} \lambda_n \, |u_n\rangle\langle u_n| \, ,$$

with the λ_n's being positive. Let $(v_n)_{n\in\mathcal{M}}$ be an orthonormal basis of $\mathcal{K}$, if we decompose Ψ in the orthonormal basis $\{u_n \otimes v_m \, ; \; n \in \mathcal{N}, m \in \mathcal{M}\}$ we get

$$\Psi = \sum_{\substack{n\in\mathcal{N}\\ m\in\mathcal{M}}} \alpha_{nm}\, u_n \otimes v_m = \sum_{n\in\mathcal{N}} u_n \otimes w_n \, ,$$

where

$$w_n = \sum_{m\in\mathcal{M}} \alpha_{nm}\, v_m \, .$$

Computing the partial trace $\mathrm{Tr}_{\mathcal{K}}\left(|\Psi\rangle\langle\Psi|\right)$ with this representation gives

$$\rho = \sum_{n,n'\in\mathcal{N}} \langle w_n \, , \, w_{n'}\rangle \; |u_n\rangle\langle u_{n'}|$$

which imposes

$$\langle w_n \, , \, w_{n'}\rangle = \delta_{n,n'} \, \lambda_n \, ,$$

if we compare with the diagonal form of ρ. Put $\widehat{w}_n = w_n/\sqrt{\lambda_n}$, for those λ_n which are strictly positive.They form an orthonormal family, which we can complete in an orthonormal family indexed by $\mathcal{N}$. We have proved that Ψ can be written as

$$\Psi = \sum_{n\in\mathcal{N}} \sqrt{\lambda_n}\, u_n \otimes \widehat{w}_n$$

for some orthonormal family $(\widehat{w}_n))_{n\in\mathcal{N}}$ in $\mathcal{K}$.

The same computation applied to Ψ' impose that it can be decomposed as

$$\Psi' = \sum_{n\in\mathcal{N}} \sqrt{\lambda_n}\, u_n \otimes \widetilde{w}_n$$

for some orthonormal family $(\widetilde{w}_n))_{n\in\mathcal{N}}$ in $\mathcal{K}$.

As the families $(\widehat{w}_n)_{n\in\mathcal{N}}$ and $(\widetilde{w}_n)_{n\in\mathcal{N}}$ are orthonormal in $\mathcal{K}$, there exist a unitary operator U on $\mathcal{K}$ such that $\widetilde{w}_n = \mathsf{U}\,\widehat{w}_n$ for all n. This proves the announced relation $\Psi' = (\mathsf{I}\otimes\mathsf{U})\Psi$. $\square$

5.3.4 Statistical Interpretation of Density Matrices

The density matrices represent the generalization of the notion of wave function which is necessary to handle open quantum systems. Their decomposition under the form (5.3) can be understood as a mixture of wave functions. It can also be understood as a random pure state.

Indeed, let $\rho = \sum_i \lambda_i \,|e_i\rangle\langle e_i|$ be a density matrix on $\mathcal{H}$. Let $|\Psi\rangle$ be a random state which can be equal to each of the $|e_i\rangle$'s with probability λ_i, respectively. We claim that when measuring any observable X on $\mathcal{H}$ with one or the other state gives the same results. Let us prove this fact.

We have seen that measuring the observable X gives a value in A with probability $\mathrm{Tr}\,(\rho\, \xi_{\mathsf{X}}(A))$. Let us detail what happens if we measure the observable X with the random state $|\Psi\rangle$. Let us adopt a helpful notation. By "(choose e_i)" we mean the event of choosing the state $|e_i\rangle$. By "(measure in A)" we mean that the event of having the result of the measure of X lying in A. Adopting obvious probabilistic notations, we know that

$$\mathbb{P}(\text{choose } e_i) = \lambda_i \qquad \text{and} \qquad \mathbb{P}(\text{measure in } A \,|\, \text{choose } e_i) = \left\| \xi_{\mathsf{X}}(A)\, |e_i\rangle \right\|^2 .$$

In particular,

$$\begin{aligned}
\mathbb{P}(\text{measure in } A) &= \sum_i \mathbb{P}(\text{measure in } A,\ \text{choose } e_i) \\
&= \sum_i \mathbb{P}(\text{measure in } A \,|\, \text{choose } e_i)\, \mathbb{P}(\text{choose } e_i) \\
&= \sum_i \lambda_i \,\langle e_i \,,\, \xi_{\mathsf{X}}(A)\, e_i\rangle \\
&= \mathrm{Tr}\,(\rho\, \xi_{\mathsf{X}}(A)) \,.
\end{aligned}$$

Hence measuring any observable X of $\mathcal{H}$ with the density matrix ρ or with the random state $|\Psi\rangle$ gives the same results, with the same probabilities. In that sense, from the physicist point of view, these two states, when measuring observables (which is what states are meant for) give exactly the same results, the same values, the same probabilities. Hence they describe the same "*state*" of the system.

Notice that this identification also works with the reduction of the wave packet, as follows. When measuring the observable X inside the set A, we have seen that the density matrix ρ has been transformed into

$$\frac{\xi_{\mathsf{X}}(A)\, \rho\, \xi_{\mathsf{X}}(A)}{\mathrm{Tr}\,(\rho\, \xi_{\mathsf{X}}(A))} \,.$$

Had we measured X with the random state $|\Psi\rangle$ we would have measured X with one the state $|e_i\rangle$. After having measured the observable X with a value in A, the state of the system is one of the pure states

$$\frac{\xi_{\mathsf{X}}(A)|e_i\rangle}{\|\xi_{\mathsf{X}}(A)|e_i\rangle\|}\,.$$

But what is exactly the probability to obtain the above pure state? We have to compute $\mathbb{P}(\text{choose } e_i \,|\, \text{measure in } A)$. This is equal to

$$\frac{\mathbb{P}(\text{measure in } A,\ \text{choose } e_i)}{\mathbb{P}(\text{measure in } A)} = \frac{\lambda_i\,\langle e_i\,,\,\xi_{\mathsf{X}}(A)\,e_i\rangle}{\mathrm{Tr}\,(\rho\,\xi_{\mathsf{X}}(A))} = \frac{\lambda_i\|\xi_{\mathsf{X}}(A)\,e_i\|^2}{\mathrm{Tr}\,(\rho\,\xi_{\mathsf{X}}(A))}\,.$$

Hence, after the measurement we end up with a random pure state again, which is equivalent (we have just discussed this point above) to having the density matrix

$$\sum_i \frac{\lambda_i\|\xi_{\mathsf{X}}(A)\,e_i\|^2}{\mathrm{Tr}\,(\rho\,\xi_{\mathsf{X}}(A))}\,\frac{\xi_{\mathsf{X}}(A)|e_i\rangle\langle e_i|\xi_{\mathsf{X}}(A)}{\|\xi_{\mathsf{X}}(A)|e_i\rangle\|^2} = \sum_i \lambda_i\,\frac{\xi_{\mathsf{X}}(A)|e_i\rangle\langle e_i|\xi_{\mathsf{X}}(A)}{\mathrm{Tr}\,(\rho\,\xi_{\mathsf{X}}(A))}\,,$$

that is, the density matrix

$$\frac{\xi_{\mathsf{X}}(A)\,\rho\,\xi_{\mathsf{X}}(A)}{\mathrm{Tr}\,(\rho\,\xi_{\mathsf{X}}(A))}\,.$$

We have proved that the two interpretations of a density matrix are equivalent from the point of view of measuring observables and from the point of view of the resulting state, after the reduction of the wave packet.

This equivalence is of course only physical, in the sense that whatever one wants to do physically with these states, they give the same results. Mathematically they are clearly different objects, one is a positive, trace 1 operator, the other one is a random pure state.

5.3.5 Faster Than Light?

Here we illustrate the role of density matrices, in particular their interpretation as "local states" via a pseudo-paradox of Quantum Mechanics. This "paradox" is an experiment which seems to show that the axioms of Quantum Mechanics are violating the locality axiom of Relativity Theory: "No information can be transmitted faster than light". Here is the idea of the experiment.

Consider two coupled systems $\mathcal{A}$ and $\mathcal{B}$ being both represented by a state space $\mathbb{C}^2$. Let $\{e_0, e_1\}$ be an orthonormal basis of $\mathbb{C}^2$. We prepare the coupled system in the state

$$\Psi = \frac{1}{\sqrt{2}}\,(e_0\otimes e_0 + e_1\otimes e_1)\,.$$

Imagine that a person, Bob, is having control on the system $\mathcal{B}$ and acts on it by performing a measurement of an observable which is diagonal along some orthonormal basis $\{f_0, f_1\}$. If the basis $\{f_0, f_1\}$ is such that all the scalar products $\langle e_i\, ,\, f_j\rangle$ are real, we have

$$(I \otimes |f_0\rangle\langle f_0|)\,\Psi = \frac{1}{\sqrt{2}}\left(\langle e_0\, ,\, f_0\rangle\, e_0 \otimes f_0 + \langle e_1\, ,\, f_0\rangle\, e_1 \otimes f_0\right) = \frac{1}{\sqrt{2}}\, f_0 \otimes f_0$$

and in the same way

$$(I \otimes |f_1\rangle\langle f_1|)\,\Psi = \frac{1}{\sqrt{2}}\, f_1 \otimes f_1 .$$

This means that, with probability 1/2 Bob will obtain the state $f_0 \otimes f_0$ and with probability 1/2 the state $f_1 \otimes f_1$.

By performing this measurement it seems that Bob has modified the $\mathcal{A}$-part of the state by forcing it to be in the f basis also. This is the starting point for imagining the faster-than-light communication. A great number of copies of the coupled system $\mathcal{A} \otimes \mathcal{B}$ are prepared, all in the same state Ψ as above. Alice goes to Andromeda with all the systems $\mathcal{A}$, Bob stays on earth with all the systems $\mathcal{B}$. They agreed that at a given fixed time Bob sends a faster-than light information to Alice in the following way. Bob chooses between two orthonormal basis f or g, both having this property to be real with respect to the basis e. He performs measurements along the chosen basis on all its $\mathcal{B}$ systems. This puts the coupled systems in the corresponding state, as described above. Immediately after that, outside the light cone of Bob's measurement, Alice performs measurements on the systems $\mathcal{A}$ in order to know if they are in the states f or g. This way she knows the information chosen by Bob (e.g. "f" or "g") faster than light!

Where is the paradox? Actually there is no paradox at all. Alice cannot figure out what Bob choose actually, in any way she may try. Imagine that Bob has chosen the basis f. The collection of state are then half in the state $f_0 \otimes f_0$ and half in the state $f_1 \otimes f_1$. If Alice measure then along the f basis she obtains f_0 with probability 1/2 and f_1 with probability 1/2. But if she measures with respect to the basis g (or any other) she then obtains g_0 with probability

$$\frac{1}{2}\,|\langle g_0\, ,\, f_0\rangle|^2 + \frac{1}{2}\,|\langle g_0\, ,\, f_1\rangle|^2 = \frac{1}{2}$$

and g_1 with probability 1/2. She gets no information at all from what has done Bob, by acting on $\mathcal{A}$ only.

Another way to understand the above experiment and conclusion is by means of density matrices. The effect of Bob's measurement transforms the pure state Ψ into the density matrix

$$\rho = \frac{1}{2}\left(|f_0 \otimes f_0\rangle\langle f_0 \otimes f_0| + |f_1 \otimes f_1\rangle\langle f_1 \otimes f_1|\right)$$

which is nothing but the operator $1/2\,\mathrm{I}$. Whatever was the choice of the bases made by Bob, Alice could see no difference afterward. Even if Bob had not changed anything to the initial state, the resulting state on $\mathcal{A}$ for Alice is the partial trace

$$\rho_A = \mathrm{Tr}_A\left(|\Psi\rangle\langle\Psi|\right) = \frac{1}{2}\left(|e_0\rangle\langle e_0| + |e_1\rangle\langle e_1|\right) = \frac{1}{2}\,\mathrm{I}\,.$$

Actually, there is no way a local action by Bob may change the local state of Alice. This assertion can be easily proved mathematically, but we do not need to develop this point here.

5.3.6 Qubits

We shall end this lecture with a tour of very useful notations and remarks concerning the simplest non trivial example of a quantum state space: the space $\mathcal{H} = \mathbb{C}^2$.

Definition 5.7. Let (e_1, e_2) be the canonical basis of $\mathbb{C}^2$. The matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are the so-called *Pauli matrices*. They are also usually denoted by $\sigma_1, \sigma_2, \sigma_3$. Together with the identity matrix

$$\mathrm{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

they form a (real) basis of the space $\mathcal{O}(\mathcal{H})$ of observables on $\mathcal{H}$. That is, any observable X on $\mathcal{H}$ can be written

$$\mathsf{X} = t\,\mathrm{I} + \sum_{i=1}^{3} x_i \sigma_i$$

for some $t, x_1, x_2, x_3 \in \mathbb{R}$. In particular note that $\mathrm{Tr}\,(\mathsf{X}) = 2t$.

Putting $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + x_3^2}$, the eigenvalues of X are $\big\{t - \|\mathbf{x}\|\,,\, t + \|\mathbf{x}\|\big\}$. In particular, any state ρ on $\mathcal{H}$ is of the form

$$\rho = \frac{1}{2}\Big(\mathrm{I} + \sum_{i=1}^{3} x_i \sigma_i\Big)$$

with $\|\mathbf{x}\| \leq 1$. The space $\mathcal{S}(\mathcal{H})$ thus identifies to $B(0,1)$, the unit ball of $\mathbb{R}^3$, with the same convex structure. Hence the pure states correspond to the points of the unit sphere S^2 of $\mathbb{R}^3$.

In other words, any state ρ is of the form

$$\rho = \frac{1}{2}\begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix}$$

with $x^2+y^2+z^2 \leq 1$ and ρ is a pure state if and only if $x^2+y^2+z^2=1$.

Let us see this directly in some different way. One can also write a pure state as:

$$\begin{aligned} \rho &= \frac{1}{2}\begin{pmatrix} 1+\cos\theta & e^{i\varphi}\sin\theta \\ e^{-i\varphi}\sin\theta & 1-\cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \cos^2(\theta/2) & e^{i\varphi}\sin(\theta/2)\cos(\theta/2) \\ e^{-i\varphi}\sin(\theta/2)\cos(\theta/2) & \sin^2(\theta/2) \end{pmatrix} \end{aligned}$$

which is the operator $|u\rangle\langle u|$ where

$$u = \begin{pmatrix} e^{i\frac{\varphi}{2}}\cos(\theta/2) \\ e^{-i\frac{\varphi}{2}}\sin(\theta/2) \end{pmatrix}.$$

This clearly describes all unitary vectors of $\mathbb{C}^2$ up to a phase factor, that is, it describes all rank one projectors $|u\rangle\langle u|$.

Notes

The general literature proposing an introduction to Quantum Mechanics is more than huge! It is very difficult to make a choice in order to guide the reader. Anyway, as this section is made for that, we propose a few references. We really appreciate the two volumes by Cohen-Tanoudji et al. [CTL92]. This enormous work is an absolute reference. It is very complete and pedagogical, it is written in a style which is very comprehensive with a mathematical culture. For a more modern reference, see [AP09].

We mentioned the very famous Orsay experiment by Aspect's team, the articles [AR81] and [AR82] are the reference articles for this experiment.

In order to write this lecture, we have also made use of the following references : [Pre04], [Bia95] and [KM98].

References

[AFP09] Gennaro Auletta, Mauro Fortunato, and Giorgio Parisi. *Quantum Mechanics.* Cambridge University Press, 2009.
[AGR81] A. Aspect, P. Grangier, and G. Roger. *Phys. Rev. Lett.*, 47, 1981.
[AGR82] A. Aspect, P. Grangier, and G. Roger. *Phys. Rev. Lett.*, 49, 1982.

[Bia95] Philippe Biane. Calcul stochastique non-commutatif. In *Lectures on probability theory (Saint-Flour, 1993)*, volume 1608 of *Lecture Notes in Math.*, pages 1–96. Springer, Berlin, 1995.

[CTDL92] Claude Cohen-Tanoudji, Bernard Diu, and Frank Laloe. *Quantum Mechanics.* 2 Volume Set. Wiley, 1992.

[KM98] B. Kümmerer and H. Maassen. Elements of quantum probability. In *Quantum probability communications*, QP-PQ, X, pages 73–100. World Sci. Publ., River Edge, NJ, 1998.

[Pre04] John Preskil. *A Course on Quantum Computation.* Course on internet: http://www.theory.caltech.edu/people/preskill/ph229/#lecture. Caltech Institute, 2004.