1

QUANTUM MECHANICS

This appendix provides a brief introduction to the axioms and language of quantum mechanics. We have no real physical pretention in this chapter; our aim is only to present the basic tools and vocabulary of quantum mechanics. We assume that the reader is familiar with the general elements of Operator Theory.

1.1 The axioms of quantum mechanics

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 precisely measured. 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 theory of 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 kind of mixture of several possible values (sometimes a continuum of possible values) which can occur with some probability. More precisely, 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 thing that 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 (such as position, energy...) of a quantum system, affects the system in an irreversible way.

An adequate mathematical language for describing the rules of quantum mechanics has been developed about 80 years ago and has shown an incredible efficiency with regards to experiments. This axiomatic language, that we

develop below, is the accepted theory nowadays by at least 99 % of the physicists. There are still researchers who work on the so-called Bohmian formalism of quantum mechanics or on the validity of the "hidden variables hypothesis", but it is not our task to compare the arguments of one or the other here. Here we place ourselves in the framework in which almost all the physicists, chemists, ... are working today and on the base of which theory the computer on which I am now typing is working.

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 rassured by 80 years of experiments agreeing with its predictions to incredible accuracy.

Let us now come to the mathematical axioms and the language of quantum mechanics.

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 φ contains all the possible information about the system; one should be able to compute any parameter of the system from φ.

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 X on $\mathcal{H}$. These are called *observables* of the system.

The set of different possible values for the measurement of an observable X is the spectrum $\sigma(X)$ of X. In particular, for some quantum systems, the energy of the system can 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 diagonalized in some orthonormal basis. This means that X can be written as

$$X = \sum_{i=1}^{n} \lambda_i P_i$$

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

$$f(X) = \sum_{i=1}^{n} f(\lambda_i) P_i.$$

Recall that, in a the infinite dimensional, an observable X is described by its spectral measure $\xi_X(\cdot)$ and by the spectral theorem:

$$X = \int_{\sigma(X)} \lambda \, d\xi_X(\lambda).$$

For every bounded measurable function f on $\mathbb{R}$ we have

$$f(X) = \int_{\sigma(X)} f(\lambda) \, d\xi_X(\lambda).$$

3rd axiom: Measurement

The only possible numerical outcome for the measurement of the an observable X is an element of its spectrum $\sigma(X)$. The result of the measurement is in the set $\sigma(X)$ but is completely random. The only possible information one can be sure of is the probability distribution of the measurement numerical 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

$$\|\xi_X(A)\,\Psi\|^2 = \langle \Psi, \xi_X(A)\,\Psi \rangle.$$

In particular, the average observed value is

$$\langle \Psi, X\Psi \rangle.$$

Furthermore, immediately after the measurement the state of the system changes to

$$\frac{\xi_X(A)\,\Psi}{\|\xi_X(A)\,\Psi\|}.$$

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

4th axiom: Dynamics

One observable of the system $\mathcal{H}$ has a particular status, the energy observable, called the *Hamiltonian* of the system. Let us denote it by H. This observable controls the way the system will evolve with time. Indeed, if one puts

$$U_t = e^{-itH}$$

for all $t \in \mathbb{R}$, then the state of the system at time t will be

$$\Psi_t = U_t \Psi_0$$

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 at time t is

$$\|\xi_X(A)\Psi_t\|^2 = \langle \Psi_t, \xi_X(A)\Psi_t\rangle = \langle U_t\,\Psi_0, \xi_X(A)U_t\,\Psi_0\rangle = \langle \Psi_0, U_t^*\,\xi_X(A)U_t\,\Psi_0\rangle.$$

But the mapping $A \mapsto U_t^*\,\xi_X(A)\,U_t$ is the spectral measure $A \mapsto \xi_{X_t}(A)$ where X_t is the observable

$$X_t = U_t^* X U_t.$$

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 states as 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 Schödinger 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) on the system as time evolves.

But when having a probabilistic point of view on this theory, as we develop all along this book, 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 distribution also evolve with time.

1.2 The Stern-Gerlach experiment

In order to illustrate the above postulates we will 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 then 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 talks 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 fixed and predictable proportions 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 this selected beam goes through a magnetic field whose direction makes an angle α with the vertical axis. One then observes another splitting into two directions (for the electrons) with the proportions $\cos^2(\alpha/2)$ and $\sin^2(\alpha/2)$ respectively.

Finally, one selects the spin $+1$ beam after this experiment (with the field in the direction α) and makes it go through the initial vertical field again. One observes a splitting into two directions, with proportions $\cos^2(\alpha/2)$ and $\sin^2(\alpha/2)$ 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 α field, the beam has lost the memory of its spin in the vertical direction.

Trying to model such an experiment and such complicated behaviour 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 modeling is clearly, at the least, very complicated, and, in fact, impossible (see next section). However, the formalism of quantum mechanics has 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 model 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

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

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

$$S_{(0,0,1)} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and $\alpha_+ = e_1$, $\alpha_- = e_2$ (the canonical basis vectors). Thus the spin of the particle is $+1$ with probability $p = |\langle \psi \, , \, \alpha_+ \rangle|^2$ and -1 with probability $q = |\langle \psi \, , \, \alpha_- \rangle|^2 = 1 - p$. 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's spin in the vertical direction again, we get probability $p = |\langle e_1 \, , \, \alpha_+ \rangle|^2 = |\langle e_1 \, , \, e_1 \rangle|^2 = 1$ to measure it with spin $+1$; and thus 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

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

The associated eigenvectors are

$$\alpha_+ = (-i\cos\theta/2, \sin\theta/2) \quad \text{and} \quad \alpha_- = (\sin\theta/2, i\cos\theta/2).$$

In particular, the respective probabilities are

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

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!

1.3 Quantum mechanics of open systems

The formalism developed in Section 1.1 needs to be extended. This formalism describes isolated quantum systems. But in many situations one has to consider quantum systems which interact with another one (or other ones). 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 ... Trying to describe the effect of such an interaction leads to a considerable extension of the axioms.

First of all, let us focus on states. A coupled system in quantum mechanics is represented by the tensor product of the corresponding Hilbert spaces $\mathcal{H}\otimes\mathcal{K}$. For the simplicity of the discussion, we suppose that $\mathcal{H}$ and $\mathcal{K}$ are both finite dimensional. A state on that system is then represented by a unit vector $\Psi \in \mathcal{H}\otimes\mathcal{K}$. This state can be decomposed as

$$\Psi = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{i,j}\, e_i \otimes f_j$$

in a typical orthonormal basis of $\mathcal{H}\otimes\mathcal{K}$.

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 M of $\mathcal{H}$, we have to consider the observable $M\otimes I$ on $\mathcal{H}\otimes\mathcal{K}$. The spectral measure of $M\otimes I$ is $\xi_M(\cdot)\otimes I$ and the probability for finding the value of the measurement lying in the Borel set A is equal to

$$\langle \Psi \, , \, (\xi_M(A) \otimes I) \, \Psi \rangle = \sum_{i,k=1}^{n} \sum_{j=1}^{m} \overline{a_{i,j}} \, a_{k,j} \, \langle e_i \, , \, \xi_M(A) \, e_k \rangle.$$

But, computing the partial trace $\rho_{\mathcal{H}}$ of the operator $|\Psi\rangle\langle\Psi|$ on $\mathcal{H}$ gives

$$\rho_{\mathcal{H}} = \mathrm{tr}_{\mathcal{K}} \, (\, |\Psi\rangle\langle\Psi| \,) = \sum_{i,k=1}^{n} \sum_{j=1}^{m} \overline{a_{i,j}} \, a_{k,j} \, |e_i\rangle\langle e_k|. \tag{1.1}$$

This means that

$$\langle \Psi \, , \, (\xi_M(A) \otimes I) \, \Psi \rangle = \mathrm{tr} \, (\rho_{\mathcal{H}} \, \xi_M(A)). \tag{1.2}$$

One can clearly see that any physical measurement that can be performed on $\mathcal{H}$ will be obtained by the same formula (1.2). One recovers the special case of an isolated system (Section 1.1) with $\rho_{\mathcal{H}} = |\Phi\rangle\langle\Phi|$:

$$\mathrm{tr} \, (|\Phi\rangle\langle\Phi| \, P) = \langle \Phi \, , \, P\Phi \rangle.$$

It is easy to check on the form (1.1) that $\rho_{\mathcal{H}}$ is always self-adjoint, positive and $\mathrm{tr} \, \rho_{\mathcal{H}} = 1$.

This means that considering open quantum systems leads to generalization of the notion of state: from now on a *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 observable X in the set A is now given by the formula

$$\mathrm{tr} \, (\rho \, \xi_X(A)) \, .$$

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

$$\rho = \sum_{n \in \mathbb{N}} \lambda_n \, |e_n\rangle\langle e_n| \tag{1.3}$$

for some orthonormal basis (e_n) of eigenvectors and some positive eigenvalues λ_n satisfying

$$\sum_{n \in \mathbb{N}} \lambda_n = 1.$$

These operators represent the generalization of the notion of wave function which is necessary to handle open quantum systems. Their decomposition under the form (1.3) can be understood as a mixture of wave functions. It can be understood as a state with a certain uncertainty: it is the wave function $|e_1\rangle\langle e_1|$ with probability λ_1,..., it is the wave function $|e_n\rangle\langle e_n|$ with probability λ_n,....

Let us see how such a generalized state should evolve with time. On the coupled system $\mathcal{H} \otimes \mathcal{K}$ we consider an Hamiltonian of the form $H = H_{\mathcal{H}} \otimes I + I \otimes H_{\mathcal{K}}$, resulting from the parallel evolution of the two systems, each one having

its own Hamiltonian: they do not interact with each other! The associated unitary group is then $U(t) = U_{\mathcal{H}}(t) \otimes U_{\mathcal{K}}(t)$ where $U_{\mathcal{H}}(t) = \exp(-itH_{\mathcal{H}})$ (and in the same way for $U_{\mathcal{K}}(t)$). Consider an initial state Ψ on $\mathcal{H} \otimes \mathcal{K}$. It gives rise to a density matrix

$$\rho_{\mathcal{H}} = \mathrm{tr}_{\mathcal{K}} \left(|\Psi\rangle\langle\Psi| \right)$$

on $\mathcal{H}$. The state Ψ evolves into the state $\Psi(t) = U(t)\Psi$ at time t. A simple computation shows that the density matrix corresponding to the state on $\mathcal{H}$ at time t is given by

$$\rho_{\mathcal{H}}(t) = \mathrm{tr}_{\mathcal{K}} \left(|\Psi(t)\rangle\langle\Psi(t)| \right) = U_{\mathcal{H}}(t)\, \rho_{\mathcal{H}}\, U_{\mathcal{H}}(t)^*.$$

This is how the evolution of states as described in Section 1.1 should be generalized when dealing with density matrices.

Even more general is the evolution of the states of $\mathcal{H}$, when the coupled system $\mathcal{H} \otimes \mathcal{K}$ undergoes a general unitary evolution. Once again we only develop the finite dimensional case here, in order to simplify the discussion.

Consider the system $\mathcal{H}$, with density matrix ρ, being coupled to the system $\mathcal{K}$ with state Ψ. Together they evolve following some unitary operator U on $\mathcal{H} \otimes \mathcal{K}$ (we forget about the time parameter t in this discussion, it may as well be supposed to be fixed):

$$\rho \otimes |\Psi\rangle\langle\Psi| \longmapsto U(\rho \otimes |\Psi\rangle\langle\Psi|)\, U^*.$$

Let us compute the partial trace of that operator on $\mathcal{H}$. If $\rho_{\mathcal{H}}$ denotes the operator $\mathrm{tr}_{\mathcal{K}}\, (U(\rho \otimes |\Psi\rangle\langle\Psi|)\, U^*)$, then it is characterized by

$$\langle a\,,\, \rho_{\mathcal{H}} b\rangle = \sum_{n \in \mathbb{N}} \langle a \otimes e_n\,,\, U(\rho \otimes |\Psi\rangle\langle\Psi|)\, U^* b \otimes e_n\rangle,$$

for all $a, b \in \mathcal{H}$ and for all orthonormal basis (e_n) of $\mathcal{K}$. For all $n \in \mathbb{N}$, consider the operator M_n on $\mathcal{H}$ defined by

$$\langle a\,,\, M_n b\rangle = \langle a \otimes e_n\,,\, Ub \otimes \Psi\rangle$$

for all $a, b \in \mathcal{H}$. Then, decomposing U as a sum of tensor products

$$U = \sum_{i,j} U_j^i \otimes V_j^i\,,$$

it is easy to check that

$$\rho_{\mathcal{H}} = \sum_n M_n\, \rho\, M_n^*, \tag{1.4}$$

which we also denote by $\mathcal{L}(\rho)$.

The operators M_n involved in (1.4) can be any kind of operators on $\mathcal{H}$, the only constraint that the unitarity of U imposes is the (easy to check) relation

$$\sum_n M_n^* M_n = I. \tag{1.5}$$

Equation (1.4), together with the constraint (1.5), describe the most general transformation that can occur to a state of a quantum system. There are called *completely positive maps* in this book, there are often known as *superoperator* in the litterature.

We stop here that discussion on completely positive maps, which is is much enough for this Appendix; it is much more developped in Chapter **??**.

1.4 Faster than light?

In this section and the next one we develop two celebrated "paradoxes" attached to the axioms of quantum mechanics.

The first one is an imaginary experience which seems to show that the axioms of quantum mechanics are violating the locality axiom of relativity theory: "No information can be trasmitted faster than light". Here is the idea of the experience.

Consider two coupled system A and 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

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

Imagine that a person, Bob, is having control on the system B and acts on it by performing a measurement along some orthonormal basis $\{f_0, f_1\}$. We have

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

and in the same way

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

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 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 transmition. A gret number of copies of the coupled system $A \otimes B$ is prepared, all in the same state Ψ as above. Alice goes to Andromeda with all the systems A, Bob stays on earth with all the systems B. They agreed that at 20:00 Bob will send a faster-than light information to Alice in the following way. Bob will choose between two basis f or g. He will perform measurements along the choosen basis on all

its B systems. This will have as effect to put the coupled systems in the corresponding state as described above. At 20:01, that is outside the light cone, Alice performs measurements on the systems A in order to know if they are in the states f or g. This way she will know the information from 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 did choose, in any way she may try. Imagine Bob has choosen 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 will obtain f_0 with probability $1/2$ and f_1 with probability $1/2$. But if she measures with respect to any other basis h (which may be g or any other) she will then obtain h_0 with probability

$$\frac{1}{2}|\langle h_0\,,\, f_0\rangle|^2 + \frac{1}{2}|\langle h_0\,,\, f_1\rangle|^2 = \frac{1}{2}$$

and h_1 with probability $1/2$. She gets no information at all from what has done Bob, by acting on 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 $1/2\, I$. Whatever was the choice of the bases made by Bob, Alice could see no difference afterwise. Even if Bob had not changed anything to the initial state, the resulting state on A for Alice is the partial trace

$$\rho_A = \mathrm{tr}_A\ (\,|\Psi\rangle\langle\Psi|\,) = \frac{1}{2}\left(|f_0\rangle\langle f_0| + |f_1\rangle\langle f_1|\right) = \frac{1}{2}I.$$

There is no way a local action by Bob may change the local state of Alice.

1.5 Hidden variables, Bell's inequalities

The second experiment that we describe is a very famous one which shows that the probabilistic behaviour 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 (in the chosen direction) respectively, 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 realisable.

In front of each particle (right and left) is placed an apparatus which measures their spin in directions $(0, \cos\alpha, \sin\alpha)$ and $(0, \cos\beta, \sin\beta)$ respectively (denoted directions α and β in the following). Let us denote by P_α the observable "spin of the left particle in the direction α", and by Q_β the observable "spin of the right particle in the direction β". We denote by

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

the corresponding eigenvectors. The probability that the spin is measured to be $+1$ on the left and $+1$ on the right is denoted by $P(P_\alpha = +1, Q_\beta = +1)$ and is equal to

$$|\langle\, \alpha_+ \otimes \beta_+ \,,\, \varphi \,\rangle|^2 .$$

A straightforward computation then gives

$$P(P_\alpha = +1, Q_\beta = +1) = \frac{1}{2}\cos^2\left(\frac{\alpha+\beta}{2}\right).$$

In the same way we find

$$\begin{aligned}
P(P_\alpha = +1, Q_\beta = -1) &= \frac{1}{2}\sin^2\left(\frac{\alpha+\beta}{2}\right) \\
P(P_\alpha = -1, Q_\beta = +1) &= \frac{1}{2}\sin^2\left(\frac{\alpha+\beta}{2}\right) \\
P(P_\alpha = -1, Q_\beta = -1) &= \frac{1}{2}\cos^2\left(\frac{\alpha+\beta}{2}\right).
\end{aligned}$$

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 ± 1-valued random variables P_α, Q_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 begining (the spin in each direction, etc...) but we have only a lack of knowledge of some of the variables of the system (this is the so-called *hidden variable* hypothesis) which leads to a random result about the quantities P_α, Q_β.

But in fact we have the following easy result.

Theorem 1.1. [Bell's 3 variable inequality] *For any three ± 1-valued random variables P_1, P_2, P_3 on a probability space $(\Omega, \mathcal{F}, P)$ we have*

$$P(P_1 = 1, P_3 = -1) \leq P(P_1 = 1, P_2 = -1) + P(P_2 = 1, P_3 = -1).$$

Proof. Simply write

$$\begin{aligned} P(P_1 = 1, &P_3 = -1) = \\ &= P(P_1 = 1, P_2 = -1, P_3 = -1) + P(P_1 = 1, P_2 = 1, P_3 = -1) \\ &\leq P(P_1 = 1, P_2 = -1) + P(P_2 = 1, P_3 = -1)\,. \end{aligned}$$ □

With the probalities we found, Bell's inequality are violated: For example take $\alpha_1 = \pi/2$, $\alpha_2 = 7\pi/6$ and $\alpha_3 = 5\pi/6$. It is impossible to attach classical random variables behind each spin of the particle. The theory and the experiment (which was perfomed in Orsay by A. Aspect's team) show the correlations we obtain cannot come from classical random variables.

A possible criticism of the above conclusion is to say that the measurement on the left polariser influences the measurement on the right one. The Orsay experiment, directed by A. Aspect in 1982, was actually a little more sophisticated. 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.

Theorem 1.2 (Bell's 4 variable inequality). *For any quadruple P_1, P_2, Q_1, Q_2 of random variables on $(\Omega, \mathcal{F}, P)$, taking the only values -1 and 1, we have*

$$P(P_1 = Q_1) \leq P(P_1 = Q_2) + P(P_2 = Q_1) + P(P_2 = Q_2)\ .$$

Proof. Note that

$$\begin{aligned} \|P_1 + Q_1\|_{L^1(\Omega)} &= \mathbb{E}\left[|P_1 + Q_1|\right] = \mathbb{E}\left[|P_1 + Q_1| \mathbb{1}_{P_1 = Q_1}\right] + \mathbb{E}\left[|P_1 + Q_1| \mathbb{1}_{P_1 \neq Q_1}\right] \\ &= 2\mathbb{E}\left[\mathbb{1}_{P_1 = Q_1}\right] \\ &= 2P(P_1 = Q_1)\ . \end{aligned}$$

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

$$\|P_1 + Q_1\|_{L^1(\Omega)} \leq \|P_1 + Q_2\|_{L^1(\Omega)} + \|P_2 + Q_2\|_{L^1(\Omega)} + \|P_2 + Q_1\|_{L^1(\Omega)}\ .$$ □

In our case, we have

$$P(P_{\alpha_i} = Q_{\beta_j}) = \cos^2\left(\frac{\alpha_i + \beta_j}{2}\right).$$

Thus the Bell inequality is violated by the choice $\alpha_1 = \pi/2$, $\alpha_2 = 7\pi/6$, $\beta_1 = 3\pi/2$ and $\beta_2 = \pi/6$.

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 timeof 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.

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.