7

The Second Mystery: Nonlocality

7.1 Introduction

In this chapter, we shall discuss the second "impossible thing" to believe before or after breakfast or the second quantum mystery: the existence of instantaneous actions at a distance in Nature. But what does "action" mean here? If there are such actions, do they allow an instantaneous transfer of matter? An instantaneous transfer of energy? An instantaneous transfer of messages? An instantaneous transfer of information? Does their existence contradict the idea that "nothing goes faster than light" (which is supposed to be a consequence of the theory of relativity)? Besides, if quantum mechanics shows that the mind acts directly on matter, do such actions at a distance justify telepathy? We have to discuss each of these points carefully and slowly.

We shall start by a little known, but very simple, thought experiment, known as Einstein's boxes. This example will allow us to raise and explain the issue of locality. Then we shall define precisely what we mean by nonlocality in Sect. 7.3 and give a simple proof in Sect. 7.4 of the fact that the world is nonlocal in a sense made explicit in Sect. 7.3.

We stress already that this proof combines *two* arguments, one due to Einstein, Podolsky, and Rosen in 1935 [65] (usually referred to by their initials EPR) and one due to Bell in 1964 [9]. We shall see over and over again that, if one considers only one of those arguments and forgets the other one, as many people do, then nothing spectacular follows. So, we shall always refer to the proof of nonlocality as the EPR-Bell result.

We should warn the reader that the views exposed here are *not* generally accepted. But we shall also try to convince the reader that this non-acceptance

J. Bricmont, *Quantum Sense and Nonsense*, DOI 10.1007/978-3-319-65271-9_7

is due to a series of misunderstandings. When it comes to the fact that the world is nonlocal, there is really no alternative!

In Sect. 7.5 we shall discuss the significance of the EPR-Bell result and some of the misunderstandings to which it gives rise.

In Sect. 7.6 we shall sketch some technological applications of quantum mechanics, in particular of the EPR-Bell result, and in Sect. 7.7 we shall discuss the tension between nonlocality and the special theory of relativity.

7.2 Einstein's Boxes

Consider the following thought experiment.[1] There is a single particle in a box B (see Fig. 7.1), and its wave function $\Psi(x)$ is non-zero everywhere in the box B. Otherwise, the precise nature of $\Psi(x)$ does not matter; one may think of a function which is constant in B for example.

One cuts the box into two half-boxes, B_1 and B_2, and the two half-boxes are then separated and sent as far apart as one wants (we assume that we can cut the box in two without affecting the particle).

According to ordinary quantum mechanics, the state becomes

$$\text{a superposition of the state } \Psi_1 \text{ and of the state } \Psi_2,$$

where the state Ψ_1 means that the particle "is" in box B_1, and the state Ψ_2 means that the particle "is" in box B_2. Here, we put scare quotes around the verb "is" because of the ambiguity inherent in the meaning of the wave function: if it reflects our knowledge of the system, then the particle *is* in one of the boxes B_i, without quotation marks. But, as we emphasized before, this is not what ordinary quantum mechanics says: it only speaks of the probability of *observing* the particle in one box or the other. If we allow for the possibility, as one should, that observations may affect the object being "observed", this distinction is crucial to make.

The state discussed here is quite similar to the one described in Figs. 4.5–4.7.

According to ordinary quantum mechanics, if one opens one of the boxes (say B_1) and one does *not* find the particle in it, one *knows* that it is in B_2.

[1] We base ourselves in this section on [134]; See that article or [98] for more details. We emphasize that this, like all "experiments" in this book, is a "thought experiment", meaning an experiment illustrating the theory, but not necessarily realized in practice. Some experiments described here are realized in laboratories, but when we will speak below of large distances between some subsystems, one should remember that we always assume implicitly that the subsystems under consideration are isolated from outside influences, a condition which is difficult to satisfy in practice if the separation between them is very large.

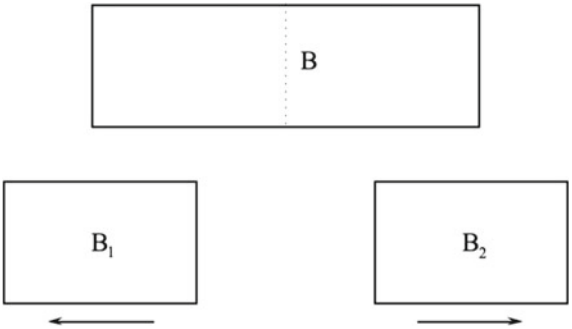

Fig. 7.1 Einstein's boxes. Reproduced with permission from T. Norsen: Einstein's boxes, *American Journal of Physics* 73, 164–176 (2005). Copyright 2005 American Association of Physics Teachers

Therefore, the state "collapses" instantaneously: it becomes Ψ_2 (and if one opens box B_2, one will find the particle in it!).

Here is the important point: since B_1 and B_2 can be as far apart as we wish, if we reject the notion of action at a distance, then it follows that acting on B_1, namely opening that box, cannot have any physical effect whatsoever on B_2. However, if opening box B_1 leads to the collapse of the wave function into one where the particle is necessarily in B_2, it must be that the particle was in B_2 all along. That is, of course, the common sense view and also the one that we would reach if the particle was replaced by any large enough object, for example a little but visible ball.

But in the situation of the particle in the box, if we reject the possibility of actions at a distance, then we must admit that quantum mechanics is *"incomplete"*, in the sense already discussed in Chap. 5: there exist other variables than the wave function that describe the system, since the wave function does not tell us which box the particle is in and we just showed, assuming no action at a distance, that the particle *is* in one of the two boxes, before one opens either of them.

In the boxes situation, the variable would simply be the label of the box in which the particle actually is. This would be an instance of what was called a "hidden variable" in Sect. 5.2.

Introducing such variables is *not* forbidden by the no hidden variables theorems of that section: the latter forbids the introduction of hidden variables for both positions and velocities of particles, but not a priori for positions alone.

This notion of action at a distance was anathema to Einstein; we saw in Sect. 4.5 that the nonlocal character of the collapse rule was already one of his objections to orthodox quantum mechanics at the Solvay Congress of 1927.

In his discussions with his colleague Max Born, Einstein wrote:

> When a system in physics extends over the parts of space A *and* B, then that which exists in B should somehow exist independently of that which exists in A. That which really exists in B should therefore not depend on what kind of measurement is carried out in part of space A; it should also be independent of whether or not any measurement at all is carried out in space A.
>
> Albert Einstein [35, p. 164]

In the example of the boxes, A and B refer here to the places where the half-boxes are (far apart) and the statement of Einstein simply means that opening one half-box cannot possibly influence the physical situation in the other half-box.

So, given his rejection of nonlocality, Einstein thought that his example of the boxes had shown that quantum mechanics is incomplete. And his reasoning was perfectly correct, if one assumes locality of course.

But if one does not reject a priori the idea of nonlocality, one should agree that Einstein had proven *at least* the following dilemma: either there exists some action at a distance in Nature (opening box B_1 changes the physical situation in B_2) *or* quantum mechanics is incomplete.

What could be nonlocal here? For example, one could think that the particle is in neither of the half-boxes before one of them is opened, and is created entirely in one of these boxes, once one of them is opened. Or, one can also think that there is one-half particle in each box and one half "jumps" instantly from one half-box to the other when one of them is opened.

This may seem extraordinarily strange (it is!), but our point here is just to indicate what seems to be an unavoidable dilemma. If you don't believe in nonlocality, then you have to accept the incompleteness of quantum mechanics.

Before discussing further this dilemma, let us consider several examples from daily life that would raise a similar dilemma and where one would side with Einstein in making assumptions, even very unnatural ones, that would preserve locality:

- Suppose that two people are located far apart, and both toss coins repeatedly but each time simultaneously. The results are completely random, heads or tails, but, at each tossing, they are always the same for both people.
- Suppose that in two casinos, far away from each other, the roulette always ends up on the red or black color, randomly, but always the same in both casinos at the same time.
- Imagine twins far apart that behave exactly in the same fashion.

In all these examples (and in many others that are easy to imagine), one would naturally assume (even if it sounded very surprising) that the two coin tossers or the casino owners were able to manipulate their apparently random results and coordinate them in advance or, for the twins, one would appeal to a strong form of genetic determinism. Who would suppose that one coin tosser immediately affects the result of the other tosser, far away, so that this other result is perfectly correlated with his own result, or that the spinning of the ball in one casino affects the motion of the ball in the other casino, again to produce a perfect correlation between both casinos, or that the action of one twin affects the behavior of the other twin? In all these cases, one would assume, even without thinking about it, a "locality" or no-action-at-a-distance hypothesis; denying it would sound even more surprising than whatever one would have to assume to explain those odd correlations.

But one thing should be a truism, namely that those correlations, if they existed, would pose a dilemma: either the results are coordinated in advance or there exists some form of action at a distance.

Note also that Einstein's assumption in the case of the boxes (that the particle is in one of the boxes before one opens either of them, which means that quantum mechanics is incomplete), is similar to the assumptions we made here about coin tossers, casinos, and twins, namely that there exists some "hidden variables" (for the coin tossers and the casinos it would be the manipulation and preparation of the results, for the twins it would be the genes) that explains the correlations. And those assumptions are very natural.

As an aside, let us mention that the example of the boxes also raises a serious question about the transition from quantum to classical physics, where "classical" just means that things have definite properties and are not in a superposed state. Indeed, if the quantum particle is replaced by a large enough object, nobody denies that the particle *is* in one of the boxes before one opens one of them. But where is the dividing line between the quantum realm and the classical one? The transition from quantum to classical physics is usually thought of as some kind of limit, like considering large masses or large energies (compared to the ones on the atomic scale); but a limit is something that one gets closer and closer to when a parameter varies, like the mass or the energy. Here, we are supposed to go from the statement "the particle is in neither of the boxes" to "the particle is in one of them, but we do not know which one". This is an "ontological" jump (meaning a radical change in what exists) and not the sort of continuous change that can be expressed by the notion of limit.

Let us now put aside the example of the boxes and ask ourselves whether there are real nonlocal effects in Nature.

The big surprise is that one can actually *prove* the existence on nonlocal effects in Nature, but not by using simply the example of Einstein's boxes. This example served only to illustrate the idea of nonlocality, which we shall now define more precisely, before proving its existence.

7.3 What Is Nonlocality?

Let us consider what kind of nonlocality or actions at a distance would be necessary, in the example of the boxes, in order to deny Einstein's conclusion about the incompleteness of quantum mechanics. So assume that the particle is in neither box, before one opens one of them. Then, opening one box, say B_1, *creates* the particle, either in B_1 or in B_2. Assume that one opens box B_1 and that one does not find the particle; therefore the particle is created in the unopened box B_2. This creation would obviously be an action at a distance; it would have the following properties:

1. *The action should be instantaneous*: opening one of the boxes creates instantly the particle in the other box.
2. *The action extends arbitrarily far*: the fact that the particle is entirely in box B_2, once we open box B_1, does not change with the distance between the boxes.
3. *The effect of that action does not decrease with the distance*: the effect is the creation of the particle in box B_2 and that effect is the same irrespective of the distance between the boxes.
4. *This effect is individuated*: suppose we have a thousand boxes, each containing one particle, and that we cut each of them into two half-boxes, then send both half-boxes far apart from each other. Then, opening one half-box will affect the state in the other half-box (coming from the cutting in two of the same box) but not in any other half-box.
5. *That action cannot be used to transmit messages*: if we open box B_1, we learn what the state becomes in box B_2, but we cannot use that to transmit a message from the place where B_1 is to the one where B_2 is. In order to transmit a message, it is enough to be able to transmit a binary signal, namely a sequence of 0's and 1's (one could use a Morse code to re-express any regular English text into such a sequence of 0's and 1's).[2] One could

[2]The Morse code is a sequence of short and long signals that allows to code any letter or punctuation mark into a sequence of such signals. One can then also code any sentence into such a sequence. If one associates, say 0 to "short" and 1 to "long", one then converts any English sentence into a sequence of 0's and 1's.

agree that 0 correspond to the particle being found in box B_1 and 1 to it not being found in box B_1 (and, thus later found in box B_2). Now, if the person who is located where box B_1 is could decide whether the particle will be found in that box or not, she could, by repeating the experiment many times (with different half-boxes coming from the splitting in two of different boxes B, each containing a single particle), send a binary signal, i.e., a message, to the person located where box B_2 is.
However, there is no known way to choose, by acting on one box, in which of the two boxes the particle will be found. Indeed, if one repeats the experiment many times with several boxes, one obtains that the particles are sometimes in B_1, sometimes in B_2, in an apparently random and uncontrollable fashion (at least that is what quantum mechanics predicts and it corresponds to what one observes in experiments). So, there is no way to use this nonlocal mechanism (assuming that it exists) to send messages.
Of course, that also implies that one cannot transfer matter or energy, because, if such a transfer was possible, then one could use it to send messages: at each instant of time, one could decide either to send or not to send instantaneously and far away a piece of matter or of energy. One could furthermore make the following convention: sending a piece of matter or of energy is associated to sending the symbol 1 and not sending it to sending the symbol 0. And, as we just saw, sending a sequence of symbols 0 and 1 is equivalent to sending a message.

Newton's theory of gravitation had also a nonlocal aspect, but we refer to Appendix 7. A for a discussion of that aspect and a comparison with the definition given here.

The Dutch physicist Hendrik Casimir explained the fundamental problem with nonlocal actions:

> If the results of experiments on free fall here in Amsterdam depended appreciably on the temperature of Mont Blanc, on the height of the Seine below Paris, and on the position of the planets, one would not get very far.
>
> Hendrik Casimir [41], quoted in [13]

Indeed, if everything was connected with everything through nonlocal actions, then science would become impossible, because, in order to test scientific theories, one always need to assume that one can isolate some systems or some variables. For example, the results of "experiments on free fall in Amsterdam" should be independent of what happens in Paris or on the Mont Blanc.

Otherwise, one would have to take into account everything that happens in the Universe in every single experiment and that would be impossible.

Therefore, because of the problems linked with nonlocality, post-Newtonian physics has tried to eliminate property 1, the instantaneity of the physical effects. For example, in the theory of electric and magnetic field, there are waves that propagate very fast, at the speed of light (about 300,000 km/s), but at a finite speed nevertheless. The same thing is true in the general theory of relativity.

One may ask whether quantum mechanics proves that there are physical effects displaying properties 1–5 above. The example of Einstein's boxes does not allow that conclusion, because one can consistently think that the quantum description is not complete and that the particle is always in one of the boxes. Indeed, that is exactly what happens in the de Broglie–Bohm theory, as we shall see in Chap. 8. In order to prove nonlocality in the sense introduced here, i.e., a phenomenon having properties 1–5 above, we have to turn to a more sophisticated situation.

7.4 A Simple Proof of Nonlocality

That more sophisticated situation is based on a two-parts argument: one part due to Einstein, Podolsky, and Rosen, in 1935, and the other one to John Bell, in 1964. Let us first explain those arguments using an analogy. The real physical situation is explained in Sect. 7.4.2. Depending on one's taste, the reader may prefer to start with the analogy or with the real thing.

7.4.1 An Anthropomorphic Thought Experiment

The analogy is with an anthropomorphic thought experiment, but which is completely similar to what happens in real experiments and could even, in principle, be realized in the anthropomorphic form presented here. Two people, which we shall call Alice and Bob (these are the habitual names used in the field of quantum information), denoted by A and B in Fig. 7.2 are together in the middle of a room and go towards two different doors, located at X and Y. At the doors, each of them is given a number, 1, 2, 3 (let's call them "questions", although they do not have any particular meaning) and has to say "Yes" or "No" (let's call that "answers").

This experiment is repeated many times, with Alice and Bob meeting together each time in the middle of the room, and the questions and answers vary apparently at random. When Alice and Bob are together in the room,

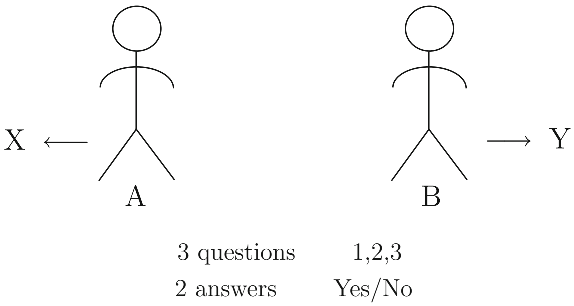

Fig. 7.2 The anthropomorphic experiment

they can decide to follow whatever "strategy" they want in order to answer the questions, namely they may coordinate their answer as they wish.

But the statistics of their answers must satisfy *two basic properties*:

1. The *first property* is that, when the same question is asked at X and Y, one always gets the same answer.
2. The *second property* is that the frequency of having the same answers on both sides when the questions are different is $\frac{1}{4}$.

How can the first property be realized? One obvious possibility is that Alice and Bob agree upon which answers they will give before moving towards the doors. They may decide, for example, that they will both say "Yes" if the question is 1, "No" if it is 2 and "Yes" if it is 3. They can choose different strategies at each repetition of the experiment and choose those strategies "at random" so that the answers will look random.

There are 3 possible questions and two possible answers to each question, so, altogether, there are 8 possible strategies:

1	2	3
Yes	*Yes*	*Yes*
Yes	*Yes*	*No*
Yes	*No*	*Yes*
Yes	*No*	*No*
No	*Yes*	*Yes*
No	*Yes*	*No*
No	*No*	*Yes*
No	*No*	*No*

Another possibility is that, when Alice reaches door X, she calls Bob and tells him which question was asked and the answer she gave. Then, of course, Bob can just give the same answer as Alice if he is asked the same question and any other answer if the question is different.

But let us assume that the answers are given simultaneously, so that the second possibility is ruled out unless there exists some instantaneous action at a distance between Alice at X and Bob at Y. Maybe Alice and Bob communicate by telepathy! Of course, this is not to be taken seriously, but that is the sort of interactions that Einstein did not consider possible. He derided them by calling them "spooky actions at a distance" [35, p. 158].

The question that the reader should ask at this point is whether there is *any other possibility*: either the answers are predetermined or a communication of some sort takes place between Alice and Bob *when* they are asked the questions. This is similar to the dilemma about the boxes: either the particle is in one of the boxes before one opens one of them, or there is some physical action between the two boxes that creates the particle in one box or the other when one opens one box. And of course it is also the same dilemma as the one for coin tossers, casinos and twins.

Note that, to pose this dilemma, one question suffices instead of three: if one question is asked in each run of the experiment, the same question at X and Y, but the answers on both sides are always the same (even though they may vary randomly between different runs of the experiment), then they must be predetermined, assuming that no communication is possible between the two sides.

Posing this dilemma is what we call the *EPR part of the argument.*

The reason that we need three possible questions is because of the *second property* of the statistics of the answers, mentioned above: when the two questions addressed to Alice and Bob are *different* (for example, question 1 is asked to Alice and question 3 is asked to Bob), then the answers must be the same *in only one quarter of the cases.*

To illustrate what we mean, we give, in Table 7.1, an example of "data": Y = Yes, N = No. These data are artificial and are meant only to give an example of what real experiments would show.

A symbol like $1N3Y$ means that the question on the left is 1 and the answer there is No, the question on the right is 3 and the answer is Yes.

When the questions on both sides are the same, the answers are always the same. They are indicated in boldface. But when the questions are different, the answers are the same only one quarter of the time. They are indicated in italics and underlined.

Table 7.1 Example of data in the "experiment" illustrated by Fig. 7.2

1 Y 1 Y	<u>*1Y3Y*</u>	1*Y*2*N*
1*N*3*Y*	2*N*3*Y*	**2 N 2 N**
1*N*2*Y*	3*Y*2*N*	1*Y*2*N*
1*Y*3*N*	**3 Y 3 Y**	**1 N 1 N**
2 Y 2 Y	<u>*1N2N*</u>	1*N*2*Y*
3*N*1*Y*	1*Y*2*N*	1*N*3*Y*
2 N 2 N	**3 N 3 N**	<u>*1Y3Y*</u>
1 N 1N	3*Y*2*N*	<u>*3N2N*</u>
1*Y*3*N*	<u>*2Y3Y*</u>	**1 Y 1 Y**
2*N*1*Y*	3*Y*2*N*	1*N*3*Y*
2 N 2 N	<u>*3N1N*</u>	**1 Y 1 Y**
<u>*2Y1Y*</u>	**1 N 1 N**	1*N*3*Y*
2*N*3*Y*	3*Y*2*N*	1*N*2*Y*
2 Y 2 Y	3*N*1*Y*	**3 Y 3 Y**
1*Y*3*N*	2*N*1*Y*	<u>*3Y2Y*</u>
1 N 1 N	1*N*2*Y*	3*Y*2*N*
<u>*2N1N*</u>	**2 N 2 N**	**1 Y 1 Y**
3 N 3 N	3*N*2*Y*	1*N*3*Y*

There are 54 results, with 18 (which is a third of the total) of them having the same questions on both sides[3] and 9 questions where the answers are the same with different questions on both sides, which is a quarter of the number of results with different questions on both sides ($9 = \frac{54-18}{4}$).

The fact that, when the questions addressed to Alice and Bob are different, the answers are the same only in one quarter of the cases, may sound innocuous.

However, this fact, combined with the idea that the properties are predetermined, leads to a contradiction:

Theorem (Bell). We cannot have these two properties together:

1. The answers are determined before the questions are asked and are the same on both sides.
2. The frequency of having the same answers on both sides when the questions are different being $\frac{1}{4}$.

Although the proof of this theorem is extremely simple, we shall defer it to Appendix 7.B.

This theorem is what we call the *Bell part of the argument.*

---

[3]If the questions are asked randomly, since there are nine possible pairs of questions, three of which are the same on both side, one expects the same questions to be asked about one third of the time.

If we combine both the EPR and the Bell parts, one has to conclude that *there are nonlocal effects in Nature*. Indeed, the EPR part shows that if there are no such effects, then the answers must be pre-determined. The frequencies of having the same answers on both sides when the questions are different being $\frac{1}{4}$ is an empirical fact. Then, Bell shows that the combination of these two statements leads to a contradiction. Since nonlocality was the *only assumption* of EPR, this assumption has to be false.

Before drawing conclusions from what has been proven, let us explain why the nonlocality proven by EPR-Bell does indeed have the properties 1–5 discussed in Sect. 7.3.

1. The effect is in principle instantaneous, but one cannot check instantaneity experimentally. However, it can at least propagate at speeds far greater than the speed of light, something that can be checked experimentally (that speed is at least 50,000 times the speed of light [90]).
2. The effect extends arbitrarily far at least in principle, that is, as long as our particles are isolated, which is difficult to realize in practice for long distances.[4]
3. The effect does not decrease with the distance between X and Y. We always get perfect correlations when the questions are the same and the same statistics for different questions.
4. The effect is individuated: if one were to send a thousand pairs of people towards the doors, one would get perfect correlations between the answers in each pair but no correlation whatsoever between the pairs.
5. Finally, this effect cannot be used to send messages from X to Y. The reason for this impossibility is similar to the one applying in the case of Einstein's boxes. Each side sees a perfectly random sequence of yes and no answers and there is no way to control what the answers will be. And, for the reasons given in Sect. 7.3, this impossibility of sending messages implies that one cannot use this mechanism in order to transfer matter or energy either.

As a historical note, let us mention that, as we said, the idea that the answers must be pre-determined if they are always the same on both sides and if there is no action at a distance is due to Einstein, Podolsky and Rosen in their 1935 paper [65]. They did not express this idea in the form used here, which is rather due to Bohm [23], but the basic idea was there.

[4]However, experiments done in 2017 by Chinese scientists show that those correlations can be maintained over more than 1,000 km [5].

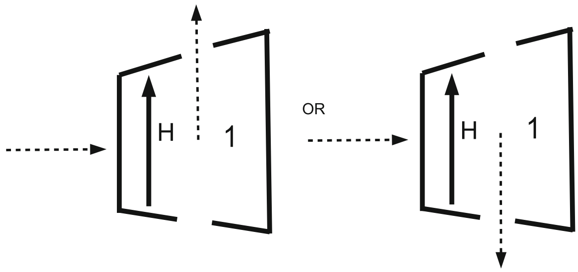

Fig. 7.3 A particle is sent towards a *box*, which is perpendicular to the plane of the figure, and in which there is a magnetic field H oriented upwards along the vertical axis, denoted 1. The particle will either go up, as on the left of the figure, viz. in the direction of the field, or down, as on the right of the figure, viz. in the direction opposite to the one of the field

Then, much later, in 1964, Bell [9] noticed that this property of the answers being pre-determined is incompatible with the frequency of having the same answers on both sides when the questions are different equal to $\frac{1}{4}$. Again, he did not formulate his argument in the form given here (which comes from [61]), but the basic idea was the same.

7.4.2 The Real Quantum Experiment

We will not discuss in detail how the quantum experiments work, but simply outline the basic idea. Particles such a electrons have a property called spin which, for our purposes, only means that if those particles are sent in a box with a magnetic field in it, they will either go in the direction of the field or in the direction opposite to the field. We will not need or use *any other notion about what "spin" means*. In particular, one should not try to "visualize" the electron as being some little particle spinning on itself.[5]

Moreover, the magnetic field can be oriented in any direction we choose and we always see the particle going in the direction of the field or in the opposite one. This is illustrated in Fig. 7.3 when the field, denoted H (a common notation for a magnetic field) is oriented vertically and in Fig. 7.4, when the field H is oriented horizontally.

One can prepare pairs of particles denoted A and B, coming from a common source, that are sent in opposite directions and have their spin measured by

[5]The actual experiments are made with photons instead electrons, with the polarization of photons replacing the spin of the electrons. It will be easier for us to discuss everything in terms of electrons and spin. Moreover, the original EPR argument (see Sect. 10.1.2) did not use spin variables. This version of the argument is due to David Bohm [23].

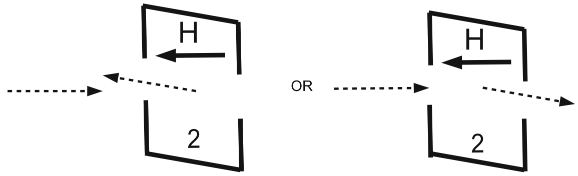

Fig. 7.4 A particle is sent towards a *box*, which is perpendicular to the plane of the figure, and in which there is a magnetic field H oriented along the horizontal axis, denoted 2. The particle will either go in the direction of the field, as on the left of the figure, or in the direction opposite to the one of the field, as on the right of the figure

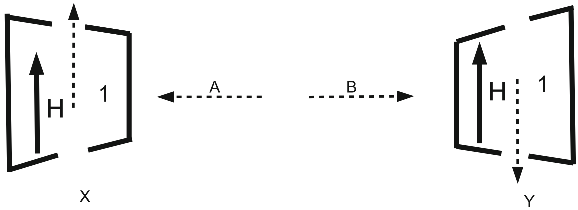

Fig. 7.5 Two particles, A and B, are sent towards *boxes*, located at X and Y, that are perpendicular to the plane of the figure, and in which there is a magnetic field H oriented upwards along the vertical axis, denoted 1. One possibility is that particle A goes up, viz. in the direction of the field, and particle B goes down, viz. in the direction opposite to the one of the field. The other possibility is shown in Fig. 7.6

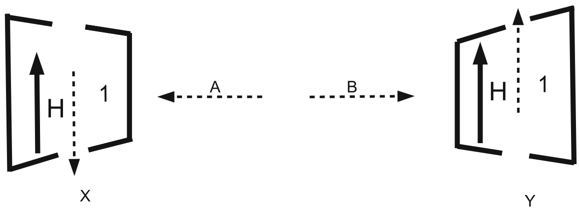

Fig. 7.6 Two particles, A and B, are sent towards *boxes*, located at X and Y, that are perpendicular to the plane of the figure, and in which there is a magnetic field H oriented upwards along the vertical axis, denoted 1. One possibility is that particle A goes down, viz. in the direction opposite to the one of the field, and particle B goes up, viz. in the direction of the field. The other possibility is shown in Fig. 7.5

the same detectors as in Figs. 7.3 and 7.4 but now put in the line of flight of each particle. The results will always be as Fig. 7.5 or 7.6: if particle A goes in the direction of the field, particle B goes in the direction opposite to the one of the field (Fig. 7.5) or vice-versa (Fig. 7.6).

We *never* see both particles going in the direction of the field or both going in the direction opposite to the one of the field.

The same thing happens when the field H is oriented horizontally, see Figs. 7.7 and 7.8.

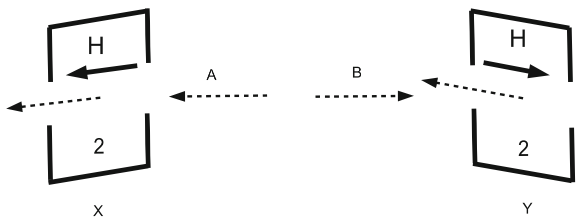

Fig. 7.7 Two particles, A and B, are sent towards *boxes*, located at X and Y, that are perpendicular to the plane of the figure, and in which there is a magnetic field H oriented along the horizontal axis, denoted 2. One possibility is that particle A goes in the direction of the field, and particle B goes in the direction opposite to the one of the field. The other possibility is shown in Fig. 7.8

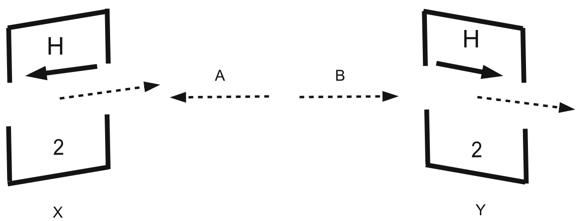

Fig. 7.8 Two particles, A and B, are sent towards *boxes*, located at X and Y, that are perpendicular to the plane of the figure, and in which there is a magnetic field H oriented along the horizontal axis, denoted 2. One possibility is that particle A goes in the direction opposite to the one of the field, and particle B goes in the direction of the field. The other possibility is shown in Fig. 7.7

We will not write down the wave function of the pair of particles in that situation, but only describe its main properties. That wave function does not assign a given value of the spin in any direction: the particle has probability one-half to go in the direction of the field and one-half to go in the direction opposite to the one of the field. But it has also the property that the two particles are correlated: if one particle goes in the direction of the field, the other particle goes in the direction opposite to the field and vice versa, each outcome having probability one-half.

It is again an instance of a superposition; the joint wave function of the pair of particles is a superposition of two wave functions: one wave function with particle A going in the direction of the field and particle B going in the direction opposite to the one of the field, and another wave function with particle A going in the direction opposite to the one of the field and particle B going the direction of the field. And that remains true for *whichever direction* one chooses to orient the magnetic field, vertical, horizontal or any other one.

Such wave functions are called *entangled*, a word introduced by Schrödinger, which reflects the fact that, no matter how far apart the particles are, they are not independent of each other.

The nature of this joint wave function means that the spin of each particle is undetermined before the measurements, if we consider only the information contained in that wave function, but the results are perfectly correlated.

This is like the two people in Sect. 7.4.1 always giving the same answer when they are asked the same question: asking a question corresponds here to measuring the spin in a certain direction and the particle going up or down corresponds to an answer yes or no. A small caveat: for the spins, we always get opposite results: in the direction of the field on one side and in the direction opposite to the one of the field on the other, while for the two people we always get the same answer. But that is just a matter of conventions: let us decide that a result in the direction of the field at X corresponds to a Yes answer there but to a No answer at Y, while a result in the direction opposite to the one of the field at X corresponds to a No answer there but to a Yes answer at Y. In that way, we will always get the same answers at X and Y, since we get opposite results for the spin.

Below, we will speak of (anti)-correlations to refer to the perfect correlations between the results at X and Y for the spin measurements.

Now we can raise the question of EPR: if there are no actions at a distance of any sort, how come the results are perfectly (anti)-correlated, no matter how far the particles are? As we discussed already, the only possibility is that the observed values of the spin, up or down, are predetermined, for each pair of particles, and for each direction.

But then comes Bell's part of the argument: one can choose three different directions in which to measure the spin (see Figs. 7.9 and 7.10), so that the answers will always be the same when the measurements are made in the same directions on each side (meaning that, if particle A goes in the direction of the field, particle B goes in the opposite direction and vice-versa), but will be the same only $\frac{1}{4}$ of the time when the measurements are made in two different directions on each side.

That is just the result of a simple quantum mechanical computation (simple, but too advanced for this book, see for example [36, p. 127] for this calculation).

The theorem of Sect. 7.4.1 then shows that this leads to a contradiction.[6]

---

[6]It is actually easy to realize the "experiment" described in Sect. 7.4.1, with Alice, Bob, and the three questions: send Alice and Bob towards X and Y and let them orient their respective magnetic field in a direction corresponding to the question that they are being asked and send towards both of them a pair of correlated particles with the quantum state described in this section. Alice and Bob can simply give an answer Yes or No depending on the result that they obtain (with our conventions), and they will then reproduce the statistics that are shown by Bell to be impossible without some form of action at a distance.

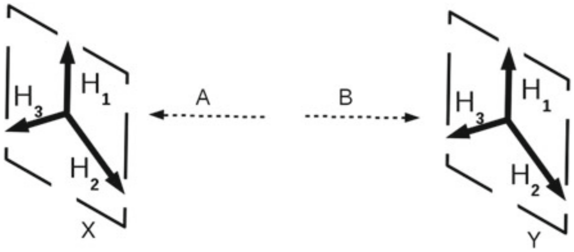

Fig. 7.9 Two particles, A and B, are sent towards *boxes*, located at X and Y, that are perpendicular to the plane of the figure, and in which there are three possible directions for the orientation of a magnetic field, denoted H_1, H_2, H_3. One repeats several times the experiment, choosing the directions of the field on each side randomly and independently of the choice on the other side. Whenever the fields are chosen in the same directions, the two particles go in opposite directions, like in Figs. 7.5, 7.6, 7.7 and 7.8

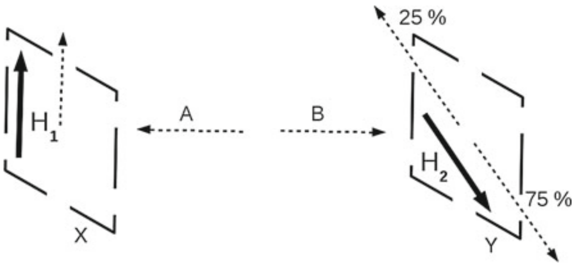

Fig. 7.10 Two particles, A and B, are sent towards *boxes*, located at X and Y, that are perpendicular to the plane of the figure, and in which there are three possible directions for the orientation of a magnetic field, denoted H_1, H_2, H_3, see Fig. 7.9. Here one chooses direction 1 at X and direction 2 at Y. In that situation, if particle A goes in the direction of the field, as in the figure, particle B will go in the direction of the field 75% of the time and in the opposite direction 25% of the time (and vice-versa). Indeed, the directions taken by particles A and B are perfectly (anti)-correlated when the directions are the same (see Figs. 7.5, 7.6, 7.7 and 7.8) but they are (anti)-correlated only 25% of the time when the directions are different. One obtains similar results for the five other possible choices of different field orientations at X and Y

Before discussing the meaning of this theorem in the next section, it is important to understand how the EPR experiment is described in ordinary quantum mechanics. As we said, the entangled wave function of the system is a superposition of two wave functions: one with particle A going in the direction of the field and particle B going in the direction opposite to the one of the field, and another one with particle A going in the direction opposite to the one of the field and particle B going in the direction of the field.

When one measures, say, the spin of particle A and that this particle goes in the direction of the field, then the superposed wave function collapses to the wave function where "particle A goes in the direction of the field and particle B goes in the direction opposite to the one of the field". Then, particle B will necessarily go in the direction opposite to the field when the spin of particle B is measured. If particle A went in the direction opposite to the field, then the collapse would occur on the other part of the entangled wave function of the system, and particle B would necessarily go in the direction of the field.

It is this collapse operation that is nonlocal in ordinary quantum mechanics. But since the meaning of both the wave function and the collapse rule are unclear in ordinary quantum mechanics, it is not clear either what this nonlocality means, and this is the source of all the ambiguities and confusions in the discussions of nonlocality, when they are carried on within the framework of ordinary quantum mechanics.

7.5 The Meaning of the EPR-Bell Argument

As we said, the argument establishing nonlocality consists of two parts: first, the EPR part shows that, if there are no pre-existing values, then the perfect (anti)-correlations when the directions are the same imply some sort of action at a distance.

The Bell part of the argument, i.e., the theorem of Sect. 7.4.1, shows that the mere assumption that there are pre-existing values leads to a contradiction when one takes into account the statistics of the results when the questions are different on both sides (or, in the real quantum experiment, when the directions in which the spin is measured are different).

The time gap between the publications of the two parts is about 30 years (1935 for EPR and 1964 for Bell). One of the problems is that many physicists react by considering only one argument, and do not connect both together (we'll come back to this in Chap. 10 and discuss how several famous physicists have reacted to the EPR-Bell argument).

If presented with the EPR argument, they will shrug their shoulders and say, well it is not surprising that the results are correlated, since the two people Alice and Bob, or the two particles, come from the same source. But that of course means that one assumes that Alice and Bob coordinate their answers in advance, or that the source prepares the particles so that the results of the various possible spin measurements are pre-determined, which is not what ordinary quantum mechanics says, since the wave function of the pair of particles does not specify those values. In other words, they implicitly agree with Einstein that locality requires the introduction of "hidden variables", namely of pre-determined spin values in the situation invented by EPR.

If presented with Bell's part of the argument they will see it as a "no hidden variables theorem", similar to the one in Sect. 5.2, meaning simply that there are no pre-existing answers or pre-existing values of the spins.

But since the orthodoxy has told us that quantum mechanics is complete, and that "hidden variable" is a bad word, they will see nothing new in that.

Bell, however, presented his result *in combination with* the EPR argument, which shows that the mere assumption of locality, combined with the perfect (anti)-correlations when the directions of measurement (or the questions) are the same, implies the existence of those hidden variables that are "impossible". So for Bell, his result, combined with the EPR argument, was not a "no hidden variables theorem", but a nonlocality theorem, the result about the impossibility of hidden variables being only one step in a two-step argument.

Of course, it is understandable that people shrink from accepting the idea of action at a distance. How can one believe that acting in some way here will affect the physical situation arbitrarily far and instantaneously?

But the argument is logically airtight and depends only on empirical data (the perfect (anti)-correlations and the $\frac{1}{4}$ factor) that have been verified in many similar experiments with more and more possible loopholes being closed.

However, we must consider and discuss some attempts that have been made to maintain that the world is local after all:

1. Some physicists say that quantum mechanics does predict both the perfect (anti)-correlations and the $\frac{1}{4}$ factor, so where is the problem? As long as one is not willing to reason beyond the quantum formalism, there is indeed no problem. But if one starts to wonder about what explains (and not simply predicts) the perfect (anti)-correlations, then one arrives at the EPR conclusions and, with Bell's argument, to the proof of nonlocality.
2. Another strategy is to maintain that the perfect correlations between the answers when the same questions are asked is simply a coincidence that does not need to be explained.
   But the whole of science can be seen as an attempt to account for correlations or empirical regularities: the theory of gravitation, for example, accounts for the regularities in the motion of planets, moons, satellites, etc. The atomic theory of matter accounts for the proportions of elements in chemical reactions. The effects of medicines account for the cure of diseases, etc. To refuse to account for correlations, without giving any particular reason for doing so, is in general a very unscientific attitude. As Bell puts it:

> You might shrug your shoulders and say 'coincidences happen all the time', or 'that's life'. Such an attitude is indeed sometimes advocated by otherwise serious people in the context of quantum philosophy. But outside that peculiar context, such an attitude would be dismissed as unscientific. The scientific attitude is that correlations cry out for explanation.
>
> John Bell [14, p. 152]

3. A variant of the "shrugging one's shoulders" argument, is to invoke a sort of "conspiracy": for example, that both Alice and Bob have an answer to only one question but that, each time the experiment is repeated, and no matter how many times it is repeated, that happens to be the question that is being asked to them. If we make that assumption, then the theorem of Sect. 7.4 cannot be derived (for the proof of the theorem to work, we need to assume pre-existing answers for at least three questions).
   This is similar to assuming that students do well on an exam, not because they have studied the course, but because they just happen to have studied precisely the answers to the very questions that they are being asked, without knowing in advance what they would be. Sometimes that may happen (students can be lucky), but it defies imagination that it could happen for all the students, all the time, and no matter how many students there are.
   The general problem with this sort of "solution" is that, no matter what the data are, one can always save one's favorite theory (here it would be the rejection of nonlocality) if one is willing to make sufficiently ad hoc assumptions. But, again, "outside that peculiar context, such an attitude would be dismissed as unscientific". Goldstein, Norsen, Tausk and Zanghì give the following example: "if you are performing a drug versus placebo clinical trial, then you have to select some group of patients to get the drug and some group of patients to get the placebo." But for that to work, you have to assume "that the method of selection is independent of whatever characteristics those patients might have that might influence how they react to the drug" [92, Note 17]. If, by accident, the people to whom the placebo is given were exactly those that are cured spontaneously, while those to whom the drug is given are so sick that the drug has little effect on them, then of course the study would be biased. And no matter how "random" the chosen sample is, this scenario always remain a logical possibility. It will only become more and more implausible as the size of the sample increases.
   The same reasoning applies to the *ad hoc* assumption that Alice and Bob have an answer to only one question, but that this question just happens to be the one that is being asked to them, and that this occurs in each repetition of the experiment. And of course, it also applies to the particles whose spin is being measured: it would be totally unreasonable to assume that the spins

of both particles are pre-determined and (anti)-correlated, but only in one direction and that this direction just happens to be the one in which their spin is measured.

If we reject such extreme forms of special pleading, nonlocality is there to stay. And refusing to face a problem is not the same thing as solving it. One thing is certain: nobody has yet proposed a genuinely local explanation for the perfect (anti)-correlations discussed here, and indeed nobody could do so, since Bell has proven that it is impossible.

To conclude, we have shown that some action at a distance does exist in Nature, but we have no idea what this action consists of. And we cannot answer that question without having a theory that goes beyond ordinary quantum mechanics. In ordinary quantum mechanics, what is nonlocal is the collapse of the wave function: if one does a spin measurement at X before doing one at Y, and one obtains, say, the up result, then the wave function of both particles are simultaneously reduced: it becomes the one where the spin of the A particle is up *and* the one of the B particle is down.

But that means that the wave function of the B particle instantaneously jumps when a measurement is made on the A particle. This looks like an action at a distance, but since the meaning of the wave function and its collapse is ambiguous in ordinary quantum mechanics, it is not clear that this is a real physical effect. But, as we have emphasized, if there are no physical effects whatsoever, then this means that we must have those predetermined values that lead to a contradiction.

7.6 Applications of Quantum Mechanics and of EPR-Bell

The French physicist Alain Aspect, who performed crucial experimental verifications of Bell's inequality, speaks of a "second quantum revolution", namely the one of "quantum information".[7] This includes quantum cryptography, which already exists and allows more secure encryptions than anything that can be done classically, quantum teleportation, as well as quantum computation, which is able in principle (but not yet in practice) to perform some calculations much faster than classical computers. We give below a short non technical introduction to each of these topics.

[7] See Aspect's introduction to the 2004 edition of [14].

7.6.1 Quantum Cryptography

We saw that one cannot use the EPR-Bell effects to send messages. However, one may use them in order to safely encode messages in a way that cannot possibly be deciphered.

First of all, how does one encode messages? To start with, as we explained in Sect. 7.3, sending a message amounts to sending a sequence of 0's and 1's.

Now, suppose that Alice and Bob are far apart and that Alice wants to send a message to Bob. She wants to be sure that no spy (usually named Eve, but one might also call it the NSA) could intercept and decipher her message. Obviously, just sending the sequence of 0's and 1's that corresponds to a Morse encoding of the message (see Sect. 7.3) won't work, because a spy could decipher the message by noticing regularities (the letter a will appear in English more often than the letter z for example, but there are many other regularities) and then simply guessing (of course, there are more sophisticated ways to do that).

What Alice and Bob need is a sequence of 0's and 1's which looks random and that is known only to themselves (see Sect. 3.1.1 for the notion of "random" sequence). This sequence is called a "key", denoted by the letter k below.

Then, once they possess such a common random sequence k, Alice and Bob can code their message into a sequence of 0's and 1's that looks random also. We explain how to do that in Appendix 7.C.

Of course, Alice and Bob still have to share this sequence k. Alice could for example toss a coin many times and count 0 if the result is heads and 1 if it is tails. That would give her a random sequence, but how to share it with Bob? If she sends it by any ordinary means of communication, it can be intercepted by a spy and the whole scheme described here would become useless.

But there is a trick based onEPR-Bell that does the job: suppose that there is a machine, situated half-way between Alice and Bob that sends to each of them one of the two particles with the wave function discussed in Sect. 7.4.2, where both particles are always (anti)-correlated when their spins are measured in the same direction. And suppose that Alice and Bob can measure the spin of those incoming particles in a given direction, but the same for both of them and chosen once and for all. Suppose further that the experiment is repeated many times.

Because of the perfect (anti)-correlations of the results when the directions in which the spin is measured are the same, Alice and Bob will have the same answer Yes or No in each measurement.[8] Then they will both share the same

[8] As we explained in Sect. 7.4.2, because of the (anti)-correlations, one makes the answer "Yes" correspond to the spin being along the direction of the field on one side and in the direction opposite to the field on the other side, and vice-versa for the answer "No".

sequence of Yes/No and they can then use that, by converting, say, each Yes into a 1 and each No into a 0, to share a common key. And, since quantum mechanics predicts that the results of the spin measurements are random, the sequence of 0's and 1's in their key will be random also.

There is still a loophole: a spy could catch the particles while they are in flight and resend them with a wave function, chosen by him, and that will produce results, when Alice and Bob do their measurements, that he can predict (and thus know). He can also arrange things so that the results look random and thus so that Alice and Bob do not notice anything strange.

But there is a way to get out of the loophole, which we will not explain in detail. Instead of measuring the spin always in the same direction, Alice and Bob can choose at each time one of the three directions considered in Sect. 7.4.2 at random. They can tell each other in which directions (1, 2, or 3) the spin measurements have been made, *without* saying what the results are. And they communicate that openly, so that a spy can listen to them, they do not care.

But then, they both know which result has been obtained on the other side when the same direction was chosen on both sides, because of the perfect (anti)-correlations, and, if the choices are perfectly random, the same direction will be chosen approximately one third of the time.[9]

So, only at the cost of more experiments, they will share a secret sequence of 0's and 1's. But, now, if the spy catches the particles in flight and re-emits them in a wave function chosen by him so that he can predict the results in one direction (say, 1), this operation will necessarily have effects on the results when the direction of measurement is not the one chosen by the spy (remember that Alice and Bob choose their directions at random, so that there is no way for the spy to know in which direction the measurements will be made).

But then, one can show that, by exchanging openly some of their results (which means that one sacrifices those results since the spy could obtain them), Alice and Bob can detect the presence of the spy.[10]

Therefore, quantum cryptography is foolproof: Alice and Bob can share a random sequence of 0's and 1's, that they may use as a key to encrypt their messages and that no spy could possibly know, without them noticing the presence of the spy.

---

[9]That is because Alice and Bob have each three possible choices of directions, so there are $3.3 = 9$ choices of pairs of directions. Three of these choices will have the same direction for both Alice and Bob and $\frac{3}{9} = \frac{1}{3}$.

[10]The statements in the last two paragraphs are based on standard quantum mechanical calculations, but justifying them would go beyond the scope of this book.

7.6.2 Quantum Teleportation

The science fiction version of teleportation, à la *Star Trek*, goes like this: you enter into a machine that copies all the information contained in your body, disintegrates it and sends all your molecules and all the information in your body to some remote place where another machine reconstitutes you in a form identical to the original one. A more fancy version has your body teleported without sending any molecule through space, but by sending only information.

Of course, such machines do not exist and nobody plans to build them in the foreseeable future. Moreover, if teleportation was made at faster than light speeds, it would flatly contradict the theory of relativity (see Sect. 7.7).

However, it is not clear that anybody would want to walk into such a machine if it existed: what if there is some malfunction and only the first half of the programme, namely your disintegration, works?

Quantum teleportation, on the other hand, does exist, but does not involve any transfer of matter or of energy. Going back to the proverbial Alice and Bob, suppose Alice possesses a particle with a certain wave function. What she can do is to make sure that Bob will have a particle with the same wave function as the one she has, after some manipulations by Alice and by Bob, but without sending Alice's particle to Bob. In fact, the only thing that she has to send to Bob is one number, chosen in the set $\{1, 2, 3, 4\}$, which she can send by an ordinary open channel, meaning at a sub-luminal speed and in such a way that a spy could discover that number without being able to know which is the wave function being transferred.

To explain how this works in more detail, one would need to use the quantum formalism and that would go beyond the scope of this book. To get a rough idea of what goes on, let us say that Alice and Bob first share an "entangled" wave function such as the one discussed in Sect. 7.4.2. Then Alice carries out a certain measurement on the system composed of the wave function she wants to send to Bob and her part of the entangled wave function, which collapses that combined wave function into one of four possible wave functions.

Because of the entanglement of the wave function shared by Alice and Bob, the measurement on Alice's side also collapses Bob's wave function, and that is where the EPR-type nonlocality enters.

Now Alice sends to Bob, by an ordinary open channel, the result of her measurement. Since there are four possibilities, she just has to send a number, 1, 2, 3, 4, each number corresponding to one of the possible results. When Bob receives that number and therefore knows the result of Alice's measurement, he acts on his own wave function in a well-defined way, depending on that result, and he is guaranteed to obtain the wave function that Alice started with, so that the wave function of Alice will have been teleported to Bob.

If a spy intercepts the open channel transmission and knows which of the numbers 1, 2, 3, 4 was sent, he is not able, with that information alone, to reconstruct the state that Alice teleports to Bob.

If all this sounds a bit mysterious, it is because we refrain from using the quantum formalism. The main point is that one uses the nonlocal aspect of the collapse of the wave function in the EPR situation, which accounts for the perfect correlations discussed in Sects. 7.4.1–7.4.2. The rest are simply local quantum mechanical operations. The fact that the information sent by Alice to Bob by an ordinary channel (one of the numbers 1, 2, 3, 4) is not enough, by itself, to teleport the wave function, shows that the nonlocal aspect of the collapse of the wave function plays an essential role in that teleportation.

Wolfgang Pauli thought that Einstein's questions were ultimately always of the same kind as "the ancient question of how many angels are able to sit on the point of a needle" [35, p. 223]. But both quantum cryptography and quantum teleportation have their origins in the EPR 1935 paper, which was regarded by many people, not only by Pauli, as "metaphysical" and "irrelevant to physics". This is just another example where the history of science shows that it may take some time before one knows whether some theoretical, or even "metaphysical", idea is useful or not.

7.6.3 Quantum Computers

Suppose that, starting from New York, you want to visit both Chicago and Los Angeles, but you want to minimize the total distance of your trip. Obviously, you will first go to Chicago and then to Los Angeles. If you start from Chicago and want to visit both New York and Los Angeles, then you should first go to New York, and then to Los Angeles, since Chicago is closer to New York than to Los Angeles.

This is a very simple instance of the "traveling salesman problem": how to visit a number of cities while minimizing the total length of your trip. It is easy to see that this is an important practical problem and not only for traveling salesmen. One may want to minimize the total length of the connections between nodes of any network, for example in a computer or in the Internet.

While the solution is obvious in the examples given above, it is not at all obvious if one wants to visit, say, the capital of every state in the Unites States. Yet, there is a "simple" solution even for that problem: make a list of all those cities in some order, compute the length between each city and the next one in that list. Then add those lengths and repeat the operation for every possible ordering of all the cities, and choose the one with the smallest total length.

The problem with that solution is that it becomes incredibly time consuming when the number of cities N is large. For ten cities, there are already more than three million lists. For twenty five cities, the number of lists is a 26-digit number (in decimal notation).[11] Even the fastest computers cannot handle problems of such length, if one were to use the "simple" method given here.

That is why an important branch of mathematics was developed in order to find algorithms (which means a mechanical method that can be implemented on a computer) that solve such problems in a "reasonable" amount of time. But even the best techniques require an amount of computing that is prohibitively large when the number of cities is large. The hope, at least for the future, is that quantum computers can reduce the amount of time needed to solve problems such as that of the traveling salesman.

A problem, simpler than the traveling salesman one, to which much attention has been devoted in quantum computing is the one of factorization of integers. Suppose you are given a number N which is a product of two prime numbers (numbers that are divisible only by 1 and by itself) and you are asked to find these numbers.[12] For example, $15 = 3 \times 5$ or $77 = 7 \times 11$. That's easy enough, and again, there is a "simple" method that solves this problem: go through the list of all prime numbers less than $\sqrt{N}$ and check if they divide N.[13]

But again, for products of large prime numbers, this method becomes terribly time consuming, and a lot of work has been devoted to finding more efficient algorithms.

There exist algorithms using quantum mechanics that, in principle, reduce spectacularly the time necessary to find the factors of a product of two prime numbers. The first "success" of this method was to factorize $15 = 3 \times 5$, which of course was not by itself a great revelation.

The way quantum mechanics enters here is via the superposition principle and interference, described in Chaps. 2–4 and illustrated by the double-slit experiment. Explaining how this works in any detail would again go beyond the scope of this book; roughly speaking, a quantum computer produces a superposition of several solutions to a problem and uses interference to select the correct one.

But this is more easily said than done and there is still a lot of work to do before quantum computers become part of our everyday reality.

[11] For a general N, the number of lists of cities is $N! = N \times N - 1 \times N - 2 \ldots 3 \times 2 \times 1$.

[12] Solving this problem has applications in classical (i.e., non quantum) cryptography, but it would go beyond the scope of this book to explain that in detail.

[13] Obviously, if $N = p \times q$, either p or q must be less than or equal to $\sqrt{N}$.

7.7 The Trouble with Relativity*

The reader who has heard of the special theory of relativity may think that the latter implies that "nothing goes faster than light". But then, doesn't the effect discussed here, instantaneous action at a distance, contradict this statement?

Unfortunately, the answer is complicated and is, in a sense, both yes and no. Moreover, to explain why this is so in detail would go beyond the scope of this book. Or, at least, we would have to include a whole long chapter explaining the theory of special relativity.[14]

The only consequence of the theory of relativity that we need to explain is the *relativity of simultaneity*. This means that, while we naively think that there is a "now" that applies to the entire universe, i.e., it seems to make sense to say that an event here and an event on the moon happen at the same time, and while this was considered true in pre-relativistic physics, it is not the case in relativistic physics.

Briefly stated, the relativity of simultaneity simply says that, if someone passes my present location, but in a moving rocket, her present and my present will be different for distant events. Certain events that occur "now" for me will occur in the future for her and vice-versa. This is illustrated in Fig. 7.11.

This sounds fantastic at first sight and to justify it, one would have to explain the whole theory of relativity, which we shall not do. But there are many experimental situations where this relativity of simultaneity (or some similar property) can be checked, the most spectacular one being the GPS: if one did not take into account such relativistic effects (both those due to the special and to the general theory of relativity), all our indications of position would be wrong and planes, for example, would crash far away from their landing strip.

All experiments in high energy laboratories must also take the relativity of simultaneity into account.

One can also imagine two twins, one of which stays on Earth, and the other one travels in a very fast spaceship, goes far from the Earth, then makes a U-turn and comes back. When he finally comes back on Earth, he will find that he is younger than the twin that didn't travel. This thought experiment has obviously never been made with real twins, but it was invented to illustrate effects that are verified with clocks traveling in airplanes. Indeed, one can send identical and very precise clocks around the Earth in airplanes flying in opposite directions and note that they are no longer synchronized after going around the Earth.

[14] See Taylor and Wheeler [186] for a rather elementary introduction to that theory and Maudlin [122] for a careful conceptual discussion.

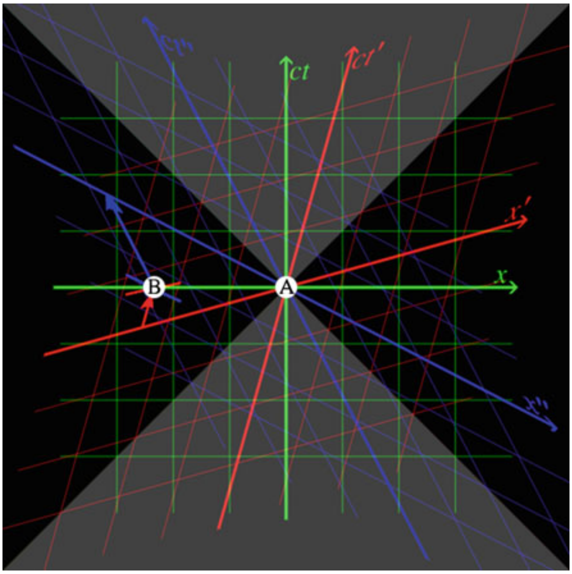

Fig. 7.11 Relativity of simultaneity: we indicate how things appear relative to three states of motion; one in *green*, one in *blue* and one in *red*. There are *green, blue* and *red lines* indicating sets of events that occur simultaneously relative to the state of motion corresponding to that color. In particular, the axis indicated x corresponds to all the events simultaneous with A relative to the state of motion indicated in *green*, the axis indicated x' corresponds to all the events simultaneous with A relative to the state of motion indicated in *red* and the axis indicated x'' corresponds to all the events simultaneous with A relative to the state of motion indicated in *blue* (the three different time axes, related to the three states of motion are denoted ct, ct', ct'', where c is the speed of light – this is a notation frequently used in relativity). So, event B is simultaneous with A relative to the state of motion indicated in *green*, but it occurred before A relative to the state of motion indicated in *blue*, and will occur after A relative to the state of motion indicated in *red* *Source* https://upload.wikimedia.org/wikipedia/commons/b/b1/RelativityofSimultaneity.svg

All these effects are not just qualitatively predicted by relativity, but are also quantitatively in very precise agreement with it.

To summarize, as surprising as the relativity of simultaneity may appear, it is a well established fact!

But what does it imply for EPR-Bell? The problem is that if simultaneity, or instantaneity, is relative to a state of motion, then with respect to which state of motion are the instantaneous actions at a distance, proven by EPR-Bell, instantaneous? Suppose that they occur simultaneously in the laboratory in which the experiment is made. Let's say that events that are simultaneous relative to the state of motion of the laboratory are represented by the green lines in Fig. 7.11, so that event A and B occur simultaneously relative to that state of motion. Relative to one state of motion different from the one of the laboratory (the blue lines in Fig. 7.11), one of these "simultaneous events" will occur *before* the other one (event B occurs before event A) and, relative to another state of motion different from the one of the laboratory (the red lines in Fig. 7.11), the same of these "simultaneous events" will occur *after* the other one (event B occurs after event A).

But that poses a serious problem for our notion of causality: indeed one would like to think that causes precede their effects in an absolute sense and one certainly would like to say that which event is a cause and which event is an effect does not depend on the state of motion relative to which those events are described.

Is there a solution to this problem? Unfortunately, not really. One possibility is to assume that there is a state of motion which is "privileged" in the sense that, relative to that state of motion, the real causes and effects occur and the causes precede their effects (for example, one could take that state of motion to be represented by the green lines in Fig. 7.11). One could consider that state of motion as one of absolute rest. This amounts to bringing back a sort of ether, which was thought, in the 19th century, to be a medium in which electromagnetic waves propagate.

The theory of relativity has not really refuted the existence of the ether, but it implies that this state of rest is not experimentally detectable, which has led to the abandonment of this notion.

Bringing back the ether does not lead to any contradiction but is somewhat unpleasant because it assumes the existence of some hidden, unobservable entity (the true state of motion in which causes and effects occur).

But giving up entirely the notion of causality is not an attractive idea either. The combination of nonlocal effects with the theory of relativity leaves us only the choice of our poison.[15]

What do "orthodox" quantum physicists say about this? In their language, as we saw in Sect. 7.4.2, it is the collapse of the wave function that is nonlocal. But the status of the wave function is ambiguous in ordinary quantum mechanics: many orthodox physicists view it as merely carrying "information" about the system, which means, if one tries to make this idea precise in our language, that particles do carry with them answers to the questions that will be asked later. And Bell has shown that this is impossible!

Alternatively, some orthodox physicists simply refuse to raise such questions, because they content themselves with "predicting results of observations", which of course they can do. But that does not remove the problem

[15] See Maudlin [122] for an elaboration of this idea.

caused by the perfect correlations and the non-existence of a local explanation of those correlations, which implies that nonlocal actions are real.

Fortunately, things could be worse but they are not. What could be worse is that, if messages could be sent instantaneously, then one could send them into one's own past! Indeed one could send instantaneously a message to a person whose state of motion is such that what is simultaneous for him lies in our past. Then, that person could re-send the message instantaneously into our own past, since our past would just be his present. This is illustrated by Fig. 7.12: If one could send instantaneously a message, then A could send a message instantaneously to B, which moves relative to A but which, at time $t = 0$ for A, is in his present. But, since B is moving relative to A, his present is not the same as the one of A. The present of B is represented by the line $t_B = 0$ in Fig. 7.12 and includes events such as A' that are in the past of A. So, if one could send instantaneously a message, then B could send the message received from A to A', that is to A in his past. In that way, A ends up sending a message to his own past.

Of course, if one could send messages into one's own past, all kinds of paradoxes would occur: you could send a message telling yourself as a student what the questions are in a certain exam, or warn yourself not to take your car

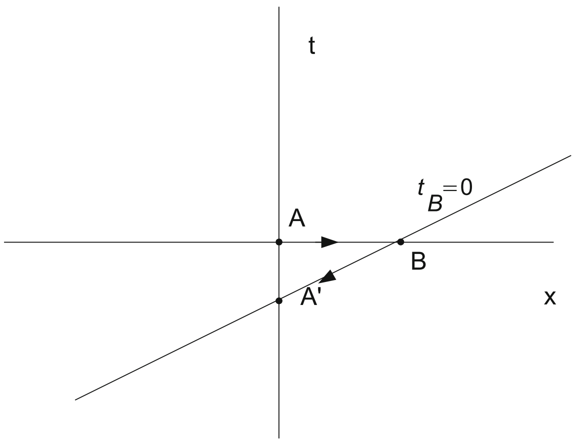

Fig. 7.12 If one could send instantaneously a message, then A could send a message instantaneously to B (indicated by the *arrow* on the x axis), which moves relative to A, but which, at time $t = 0$ (i.e., along the x axis) is in his present. But the present of B is represented by the *line* $t_B = 0$ and includes events such as A' that are in the past of A. The *line* $t_B = 0$ corresponds to one of the *red lines* in Fig. 7.11, while the axis $t = 0$ (the x axis) corresponds to a *green line* in Fig. 7.11. So, if one could send instantaneously messages, then B could send the message received from A to A' (indicated by the *arrow* on the $t_B = 0$ line), that is to A in his past. In that way, A ends up sending a message to his own past

on the day that you had an accident. Physics allow for many paradoxes and counterintuitive statements, but it cannot go that far![16]

So, to summarize: the answer to the question of the tension between relativity and nonlocality is both "no", if one means that this could allow the sending of messages into one's own past, and "yes": there is a serious problem if one wants to reconcile both ideas with a "causal" view of the world, where there are causes and effects and where the former precede the latter in a sense that is independent of the state of motion relative to which those causes and effects are described. In fact this is a major open problem in physics, although one that is not recognized as such by most physicists, because of their refusal to discuss the meaning of the wave function beyond being a tool to predict results of measurements.

7.8 Summary

In his chapter, we discussed the second and deepest mystery of quantum mechanics: nonlocality. Briefly put, nonlocality means that there are correlations between distant events that cannot be explained by antecedent causes. This implies that there must be some form of actions at a distance between the places where those correlated events occur.

We first illustrated the problem with the example of Einstein's boxes. Take a single particle in a box that is cut in two and let each half-box be sent far away from the other. When one opens one of the half-boxes, one always finds the particle in one of them. Then, one faces the following dilemma: either the particle is in one of the half-boxes before one opens one of them, but then the pure quantum mechanical description through the wave function is incomplete, or the particle is somehow created in one of the half-boxes when one opens one of them, but then some action at a distance takes place.

This action at a distance has to be instantaneous, its effect extends arbitrarily far, does not decrease with the distance between the two half-boxes, and is individuated in the sense that the action takes place only between the two half-boxes that came from the same box cut in two and containing one particle initially. Moreover, because of the random nature of the results, one cannot use this action (if it exists) to transmit messages.

[16]As we explained in Sect. 7.3 the impossibility of sending a message instantaneously implies the impossibility of sending instantaneously either matter or energy.

Because of all those properties, the actions at a distance envisioned here are radically different from the ones of Newtonian physics discussed in Appendix 7.A.

However, one cannot solve the dilemma posed by the boxes in favor of the existence of actions at a distance. Indeed, one can constantly think that the particle is in one of the half-boxes before one opens one of them; we explain that in Chap. 8.

But the existence of actions at a distance can be proven directly as we saw in Sect. 7.4: one can devise an experiment, where two people who are far apart respond to three questions in a perfectly correlated way when the questions are the same. This again poses the dilemma: either the answers are coordinated in advance or there is some action at a distance between the two people. We call this dilemma the EPR part of the argument.

But the statistics of the answers when the questions are different (only one quarter of the same responses) rule out the possibility of such a coordination. That is the Bell part of the argument. So, we are left with only one possibility: actions at a distance!

At this stage, we do not know how such actions work or what causes them. But we did show that some escape strategies that have been proposed to maintain that the world is local after all do not work.

Moreover, quantum mechanics has applications in cryptography and teleportation of information that do depend on the nonlocal aspects of Nature revealed by the EPR-Bell reasoning.

Finally, we addressed the subtle issue of the tension between quantum nonlocality and relativity. The short answer is that, because of relativity, there is a tension between causality and quantum nonlocality, but not a sharp contradiction since that nonlocality does not allow the transmission of messages.

Appendices

7.A Nonlocality in Newton's Theory

Newton's theory of gravitation has also a nonlocal aspect, but which is different from the one discussed in Sect. 7.3.

Two of the most famous laws of classical physics are due to Newton. Suppose we have two bodies, labelled 1 and 2 whose masses are denoted M_1 and M_2. Then, they attract each other through a gravitational force proportional to the product $M_1 \times M_2$ of their masses and proportional to the inverse of the square of their distance d. The second law is that "the force is equal to the mass times the acceleration", the acceleration being the rate at which the velocity changes. So, if we change the position of body 2, we change the value of the distance d between the two bodies in that equation, and therefore the value of the acceleration of body 1. If the acceleration changes, this changes the velocity of that body, and if one changes its velocity, one changes the position of the body.[17]

This makes actions at a distance possible: since the gravitational force depends on the distribution of matter in the Universe, changing that distribution, say by waving my arm, instantaneously affects the motion of all other bodies in the Universe (of course, the effect is minuscule, but we give this simple example to illustrate the principle). That action at a distance has properties 1 and 2 of Sect. 7.3, since it is instantaneous and acts arbitrarily far.

But it does not have the other properties, 3, 4 and 5 of Sect. 7.3: it does not have property 3 because its effect decreases with the distance, since the effect is proportional to the inverse of the square of the distance d. Besides, it affects all bodies at a given distance equally: there is nothing special in body 2 except its distance with respect to body 1. In other words, unlike what happens with the boxes, it is *not* individuated, so it does not have property 4.

On the other hand, it can in principle be used to transmit messages (so it does not have property 5 of Sect. 7.3): if I decide to choose, at every minute, to wave my arm or not to wave it, one can use that choice of movements to encode a sequence of zeros and ones and, assuming that the gravitational effect due to that movement can be detected, one can transmit a message instantaneously and arbitrarily far (of course, the further away one tries to transmit it, the harder the detection).

---

[17] In formulas, the first law stated here says that

$$F = \frac{GM_1M_2}{d^2} \tag{7.1}$$

where F is the force exerted by one body on the other, G is a constant (called Newton's constant), which will not concern us here, and d is the distance between the two bodies. In reality, the force is a vector and we are indicating here only its length. The same is true for the acceleration below. This law is often called the "inverse square law", because of the factor d^2 in the denominator of the right hand side of (7.1).

The second law is:

$$F = Ma, \tag{7.2}$$

where F is the force exerted on the body of mass M, and a is the acceleration of that body.

Now, let us see what this implies for the acceleration of the body 1. We need to put an index 1 in (7.1) and in (7.2): $F_1 = \frac{GM_1M_2}{d^2}$ and $F_1 = M_1a_1$, where F_1 is the force exerted on body 1, of mass M_1, and a_1 is its acceleration. Inserting the first of these formulas into the second one, and dividing both sides by M_1, we get:

$$a_1 = \frac{GM_2}{d^2}, \tag{7.3}$$

namely the acceleration of body 1 depends instantaneously on its distance with respect to body 2.

Note that all this refers to Newton's *theory*. There have been no experiments performed in this framework that could *prove* that gravitational forces really act instantaneously or at least at speeds faster than the speed of light, and this is a major difference with respect to the situation in quantum mechanics.

7.B Proof of Bell's Theorem in Sect. 7.4

As we said in Sect. 7.4.1, there are three questions numbered 1, 2, and 3, and two answers Yes and No. If the answers are given in advance, there are $2^3 = 8$ possibilities:

1	2	3
Yes	*Yes*	*Yes*
Yes	*Yes*	*No*
Yes	*No*	*Yes*
Yes	*No*	*No*
No	*Yes*	*Yes*
No	*Yes*	*No*
No	*No*	*Yes*
No	*No*	*No*

So Alice and Bob could agree, for example to always answer "Yes" to the first question, "No" to the second one and also "No" to the third (let's call that the YNN strategy). Or they could follow each of the strategies YYN, NYN, and NNN one third of the time. Or they could choose their answers at random among the eight possibilities.

In any case, in *each situation* there are at least *two questions* with the same answer. Therefore,

$$\begin{aligned} &\text{Frequency (answer to 1 } = \text{answer to 2)} \\ &+ \text{Frequency (answer to 2 } = \text{answer to 3)} \\ &+ \text{Frequency (answer to 3 } = \text{answer to 1)} \geq 1 \ , \end{aligned} \tag{7.4}$$

since at least one of the identities: answer to 1 = answer to 2, answer to 2 = answer to 3, answer to 3 = answer to 1, holds in every run of the experiment.

But if

$$\begin{aligned} &\text{Frequency (answer to 1 } = \text{answer to 2)} \\ &= \text{Frequency (answer to 2 } = \text{answer to 3)} \\ &= \text{Frequency (answer to 3 } = \text{answer to 1)} = 1/4 \ , \end{aligned}$$

we get $\frac{3}{4} \geq 1$, which is a contradiction. ■

The inequality like (7.4), saying that a sum of frequencies are greater than or equal to 1, is an example of a *Bell inequality*, i.e., an inequality which is a logical consequence of the assumption of pre-existing values, but which is violated by quantum predictions.

7.C How to Encode Secret Messages?

Suppose that Alice and Bob possess a common key k, namely a random sequence of 0's and 1's, and that they want to use that sequence in order to encode a message m which is also a sequence of 0's and 1's, but a non random one (messages that have a meaning for us are not random) in such a way that the result looks random.

They can do that by adding "in binary addition" the message to be sent m and the sequence k; binary addition means that one adds each of the corresponding symbols in the two sequences according to the following rules:

- $0 + 0 = 0$,
- $0 + 1 = 1$,
- $1 + 0 = 1$,
- $1 + 1 = 0$.

It looks like those rules are the ordinary ones except for the last one where we seemingly made a mistake: $1+1 = 2$! But these are just rules that we define to be the valid ones if we deal with only two numbers, 0 and 1. There is no symbol 2 here, by definition.

For example, if the message to be sent is $m = 01101010$ and the common sequence or key is $k = 11011001$, we have $m + k = 10110010$ and that is the sequence which is sent by Alice to Bob through an open channel (and can in principle be obtained by the spy).

Since Bob also has the sequence k, it is very easy for him to obtain the original message m. Indeed, with our binary rules, we have $0 + 0 = 0$ and $1+1 = 0$. So, adding twice the same message amounts to adding 0 everywhere, which means that one does not change anything. So, one has $m + k + k = m$. Just to check, add $k = 11011001$ to $m + k = 10110010$, with the binary rules, and you will obtain back $m = 01101010$.

Now, if the sequence k is sufficiently long and sufficiently random one can use it to encode any given message (we just gave one with eight symbols to illustrate the method) and the regularities in the message m will disappear because of the randomness in k.

To check that the sequence $m + k$ will be as random as k, no matter what m is, just consider the least random m one can imagine: $m = 111111111\ldots$ that is every symbol of m is 1. Then, with the rules of binary addition, in the sequence $m + k$ one will simply have each symbol 0 in k replaced by 1 and each symbol 1 in k replaced by 0. But if the sequence k is random, the new sequence $m + k$ will also be random.