History of Quantum Mechanics or the Comedy of Errors[1]

Jean BRICMONT

IRMP, Université catholique de Louvain, chemin du Cyclotron 2,
1348 Louvain-la-Neuve, Belgique

Abstract

The goal of this paper is to explain how the views of Albert Einstein, John Bell and others, about nonlocality and the conceptual issues raised by quantum mechanics, have been rather systematically misunderstood by the majority of physicists.

1   Introduction

The history of quantum mechanics, as told in general to students, is like a third rate American movie: there are the good guys and the bad guys, and the good guys won.

The good guys are those associated with the "Copenhagen" school, Bohr, Heisenberg, Pauli, Jordan, Born, von Neumann among others. The bad guys are their critics, mostly Einstein and Schrödinger and, sometimes de Broglie. The bad guys, so the story goes, were unwilling to accept the radical novelty of quantum mechanics, either its intrinsic indeterminism or the essential role of the observer in the laws of physics that quantum mechanics implies. They were hinged to a classical world-view, because of their philosophical prejudices.

Einstein invented various *gedanken experiments* in order to show that the Heisenberg uncertainty relations could be violated (those relations put limits on what can be known about physical systems), but Bohr successfully answered those objections. After the war, in 1952, David Bohm rediscovered an old idea of de Broglie and tried to develop an alternative to quantum mechanics that would restaure determinism, but his theory got very little attention among physicists, mostly because, by then, the successes of quantum mechanics had been so spectacular that hardly anybody had still doubts about its correctness.

---

[1] Talk given on November 11, 2014, at the conference: La mécanique à la lumière de son histoire, de la Modernité à l'Époque Contemporaine (XVIIe-XXe s.), given in honor of Prof. Patricia Radelet -De Grave.

Bohm's theory was introducing "hidden variables", meaning some unobservable quantities (hence, "metaphysical") that are not part of the standard quantum mechanical description.[2] Einstein had also favored the introduction of such variables in order to "save determinism". The first objection to this idea is obvious: why bother with unobservable entities in order to satisfy a philosophical prejudice?

However, the definite blow against hidden variables was given in 1964 by John Bell who showed that, merely imagining that such variables exist leads to predictions that are in contradiction with those of quantum mechanics. Moreover, those specific predictions were later tested in laboratories and, of course, the observations came definitely on the side of quantum mechanics and against hidden variables. Case closed!

The goal of this paper is to show that all of the above is essentially false.[3] And in order to show that, we will have to do almost no physics, just reading what Einstein and Bell really said (as Patricia always told me: "go read the sources!"). One may not agree with them, one may adhere to the standard views on quantum mechanics (whatever they are, which nowadays is not so simple to say), but one should, it seems to me, at least try to understand *what Einstein and Bell said before rejecting their ideas*!

We shall first explain what really worried Einstein about quantum mechanics, namely its nonlocal aspect.[4] Next, we will discuss what Einstein (together with Podolsky and Rosen) and Bell said about that issue of nonlocality. Then we will review how this was misunderstood by most physicists.

---

[2] Throughout this paper, we will use the expression "hidden variables" to mean any variables that are used to describe the state of a quantum system and that are not included in the usual quantum description, via a wave function or a quantum state (the latter being simply a wave function "times" some internal states such as spin states, but we will not stress that distinction here and we will speak only of wave functions).

[3] The material for this paper is in part taken from [20, 21], and we refer to those texts for more details. In particular, we will not discuss much Bohm's theory here, but only note that it is not true that this theory has been refuted Bell; see [14, 29, 30, 46, 20] for an exposition of that theory.

[4] At least that was one of his worries. He also worried about the centrality of the "observer" in quantum mechanics, but we will not discuss that problem here (see [20]). Let us nevertheless mention that this worry is illustrated by the following remark by Abraham Pais about his conversations with Einstein [58, p. 907]: "We often discussed his notions on objective reality. I recall that during one walk Einstein suddenly stopped, turned to me and asked whether I really believed that the moon exists only when I look at it." "Is the moon there when nobody looks?" has also become a famous Einstein quote (frequently misunderstood, since Einstein had no doubt that the answer to that question was "yes", while quantum mechanics is often interpreted as implying that we don't know anything about the world before we "look").

2 Einstein's Real Worries

The most often quoted phrase of Einstein is probably "God does not play dice".[5] But this is misleading. It is true that Einstein did express concerns about determinism, but that determinism was Einstein's main concern was strongly denied even by Pauli, who was, in general, on the Copenhagen side of the arguments (if one wants to use this somewhat reductionist terminology) and who wrote in 1954 to Max Born:

> Einstein does not consider the concept of "determinism" to be as fundamental as it is frequently held to be (as he told me emphatically many times) [...] he disputes that he uses as a criterion for the admissibility of a theory the question: "Is it rigorously deterministic?"

> Wolfgang Pauli [19, p. 221]

If determinism was not Einstein's main concern, what was? A clue to the answer is given in a letter, written in 1942, which also speaks of God not playing dice:

> It seems hard to sneak a look at God's cards. But that he plays dice and uses "telepathic" methods (as the present quantum theory requires of him) is something that I cannot believe for a single moment.

> Albert Einstein [36] quoted in [27, p. 68]

What Einstein called "telepathic" here is related to what nowadays is called non-locality and this is the aspect of quantum mechanics that bothered Einstein. He and Schrödinger saw clearly that aspect when everybody else had their heads in the sand, so to speak.

To understand nonlocality, let us think for a moment about the concept of wave function and what it means to say (as is usually done) that it provides a *complete description* of quantum systems. The wave function, let's say for one particle, is a (complex valued) function defined on $\mathbb{R}^3$: $\Psi(x) \in \mathbb{C}$, $x \in \mathbb{R}^3$. Its meaning, in quantum mechanics textbooks is that the square of its absolute value, $|\Psi(x)|^2$, gives the probability density of *finding* the particle at a given point, if one "measures" the particle's position. A naive interpretation of the word measurement would lead us to think that, if the particle is found somewhere, it is because *it was there* in the first place! But that is not what quantum mechanics says: indeed, "measurements" might affect the object being measured and not simply reveal any of its properties, so that the word measurement is used in quantum mechanics in a rather awkward way.

---

[5] The complete quote comes from a letter to Max Born in 1926 [19, p. 91]: "Quantum mechanics is very worthy of regard. But an inner voice tells me that this is not yet the right track. The theory yields much, but it hardly brings us closer to the Old One's secrets. I, in any case, am convinced that He does not play dice." Of course, as Einstein emphasized several times, his "God" had nothing to do with the personal gods of the "revealed" religions.

But what worried Einstein is that, after the measurement of the position, the wave function changes and collapses to a wave function concentrated at or around the point where the particle is found. But that means that the value of the wave function suddenly jumps to zero everywhere, except where the particle is found and that effect is supposed to be instantaneous. Since the wave function can be, before the measurement, different from zero far away from the point of detection, it means that the collapse of the wave function is a sort of nonlocal action, or an action at a distance, and this is what Einstein could not accept (in part, because it contradicted the theory of relativity, as he and most people understood it).[6] That action at a distance is what "telepathy" means in the above quote;[7] Einstein also called such actions at a distance "spooky" ([19, p. 158]).

Already at the 1927 Solvay Conference, Einstein expressed very clearly this worry. He considered a particle going through a hole, as shown in Fig. 1. In the situation described in the picture, the wave function spreads itself over the half circle, but one always detects the particle at a given point, on the hemispherical photographic film denoted by P in Fig. 1. If the particle is *not* localized anywhere before its detection (think of it as a sort of "cloud", as spread out as the wave function), then it must condense itself at a point, in a nonlocal fashion, since the part of the particle that is far away from the detection point must "jump" there instantaneously. If, on the other hand, the particle *is* localized somewhere before its detection, then quantum mechanics *is* an incomplete description of reality, since quantum mechanics does not include the position of the particle in its description.

Einstein contrasted two conceptions of what the wave function could mean:

> Conception I. The de Broglie–Schrödinger waves do not correspond to a single electron, but to a cloud of electrons extended in space. The theory gives no information about individual processes, but only about the ensemble of an infinity of elementary processes.

> Conception II. The theory claims to be a complete theory of individual processes. Each particle directed towards the screen, as far as can be determined by its position and speed, is described by a packet of de Broglie–Schrödinger waves of short wavelength and small angular width. This wave packet is diffracted and, after diffraction, partly reaches the screen P in a state of resolution.[8]

<div align="right">Albert Einstein [2, p. 441]</div>

---

[6] That is because the relativity of simultaneity, which is part of the theory of relativity, implies that, if causal effects can be instantaneous in some reference frames, then they must act backward in time in some other reference frames. This, to put it mildly, contradicts our intuitive notion of causality. This problem is discussed in detail in [51].

[7] To be precise, Einstein was probably referring there to the stronger sort of nonlocality implied by the EPR argument discussed in Sect. 6.

[8] That last expression is not very clear, but the original German text has been lost and the French translation [61], from which this is translated into English, is not clear either. (Note by J.B.)
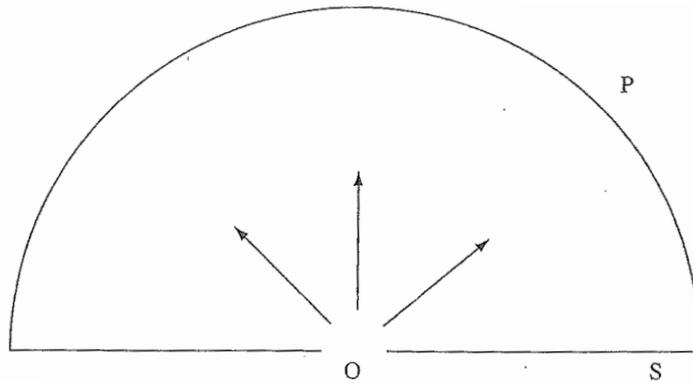
Fig. 1: Einstein's objection at the 1927 Solvay Conference. Reproduced with the permission of the Solvay Institutes and of Cambridge University Press. See [2, p. 440], or [61, p. 254] for the "original" (published in French translation at the time of the Solvay Conference). In fact Einstein raised a similar issue as early as 1909, at a meeting in Salzburg [2, p. 198]

The distinction made here was repeated by Einstein all his life. Either one adopts a statistical view (conception I) and a more complete theory can be envisioned (for "individual processes"), or one declares the quantum description "complete", but this means that the particle is spread out in space before being detected on the screen.

After explaining the merits of conception II, Einstein raises the following objection:

> But the interpretation, according to which $|\psi|^2$ expresses the probability that *this* particle is found at a given point, assumes an entirely peculiar mechanism of action at a distance, which prevents the wave continuously distributed in space from producing an action in *two* places on the screen.

Albert Einstein [2, p. 441]

Indeed, if the particle is spread out in space before being detected (which is what the expression "complete description" *means*), then the fact that it is always detected at a given point implies that it condenses itself on that point and that its presence vanishes elsewhere. Thus something nonlocal must be taking place. Einstein adds:

> In my opinion, one can remove this objection [action at a distance] only in the following way, that one does not describe the process solely by the Schrödinger wave, but that at the same time one localises the particle during the propagation. I think that Mr de Broglie is right to search in this direction.[9] If one works solely with the Schrödinger waves, interpretation II of $|\psi|^2$ implies to my mind a contradiction with the postulate of relativity.

---

[9] Einstein refers here to de Broglie's guiding wave theory, which was also presented at the 1927 Solvay Conference, see [2] or [20, chap. 8] for a further discussion of that Conference. The de Broglie's guiding wave theory was later developed by Bohm.

Thus, the essence of Einstein's objection to the orthodox view of quantum mechanics (including the conflict between nonlocality and relativity) was already expressed in 1927!

It is interesting to read Bohr's response:

> I feel myself in a very difficult position because I don't understand what precisely is the point which Einstein wants to [make]. No doubt it is my fault.

Einstein's objections became sharper as time went on, specially with the 1935 Einstein, Podolsky, and Rosen (or EPR) paper [33].

Before turning to the EPR argument, we will 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 will define precisely what we mean by nonlocality and give a simple derivation of Bell's argument (due to [28]), *combined* with the EPR argument. That is the simplest and clearest way to arrive at the main conclusion, namely that the world is nonlocal in a sense made explicit in Sect. 4.

3 Einstein's Boxes

Consider the following thought experiment.[10] There is a single particle in a box B (see Fig. 2), and its wave function is $|B\rangle$, meaning that its wave function is distributed[11] over B. 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.

According to ordinary quantum mechanics, the wave function becomes

$$\frac{1}{\sqrt{2}}\big(|B_1\rangle + |B_2\rangle\big)\,,$$

where the wave function $|B_i\rangle$ means that the particle "is" in box B_i, $i = 1, 2$ (and again, each wave function $|B_i\rangle$ is distributed in the corresponding box).[12] Here, we put

---

[10] We base ourselves in this section on [57], where the description of the experiment is due to de Broglie [25, 26]. Einstein's original idea was expressed in a letter to Schrödinger, written on 19 June 1935, soon after the EPR paper was published [40, p. 35]. However, Einstein formulated his "boxes" argument in terms of macroscopic objects (small balls) and then gave a slightly different and genuinely quantum mechanical example, but illustrating the same point as the one made here. Figure 2 is taken from [57].

[11] The precise distribution does not matter, provided it is spread over B_1 and B_2 defined below; it could be distributed according to the square of the ground state wave function of a particle in the box B.

[12] This is of course a thought experiment. We assume that we can cut the box in two without affecting the particle. But similar experiments have been made with photons and were already suggested by Heisenberg in 1929, see [49, p. 39].
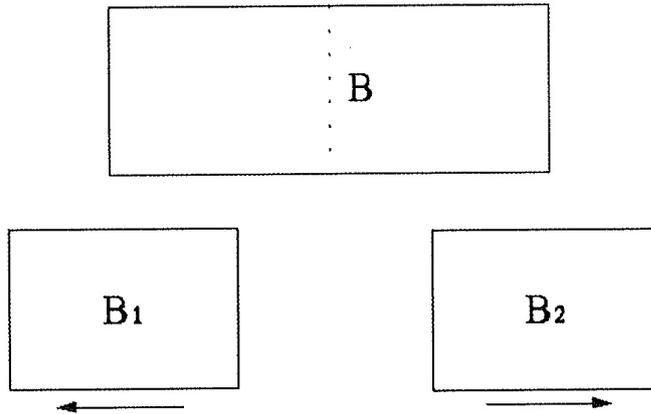
Fig. 2: Einstein's boxes. Reproduced with permission from T. Norsen: Einstein's boxes, *American Journal of Physics* **73**, 164–176 (2005). Copyright 2005 American Association of Physics Teachers

scare quotes around the verb "is" because of the ambiguity inherent in the meaning of the wave function, reflected in the two notions discussed by Einstein at the 1927 Solvay Conference: if the wave function reflects our knowledge of the system, then the particle *is* in one of the boxes B_i, without quotation marks. But if one thinks of the wave function as being physical and of the position of the particle as being created or realized only when someone measures it, then the quotation marks are necessary and "is" means: "would be found in box B_i after a measurement of its position".

Now, 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. Therefore, the wave function "collapses" instantaneously: it becomes $|B_2\rangle$ (and if one opens box B_2, one will find the particle in it!).

Since B_1 and B_2 are spatially separated (and 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.

But in the situation of the particle in the box, if we reject action at a distance, then we must admit that quantum mechanics is not complete, in the sense that Einstein gave to that word: there exist other variables (called "hidden variables") than the wave function that describe the system, since the wave function does not tell us in which box the particle is 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 any case, with this argument, Einstein had already proven 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. Since action at a distance was anathema to him (and probably to everybody else at that time), he thought that he had shown that quantum mechanics is incomplete.

There are many examples at a macroscopic level 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 each tosses coins with results that are always either heads or tails, randomly, but always the same for both throwers. Or 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. Or imagine twins far apart that behave exactly in the same fashion.[13] 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 throwers 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 one, far away, or that the spinning of the ball in one casino affects the motion of the other ball, or that the action of one twin affects the behavior of the other twin? In all these cases, one would assume a locality 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 pose a dilemma: the results are either coordinated in advance or there exists a form of action at a distance. Therefore, Einstein's conclusion in the case of the boxes (incompleteness of quantum mechanics), which is similar to the assumptions we made here about coin throwers, casinos, and twins, was actually 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. Indeed, if the quantum particle is replaced by a "classical" one, meaning a large enough object, nobody denies that the particle *is* in one of the boxes before one opens either box. 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. 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 and not the sort of continuous change that can be expressed by the notion of limit.

Let us now define more precisely what we mean by nonlocality.

4   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 (which is what quantum mechanics being complete means). Then opening one box, say B_1, creates the particle, either in B_1 or in B_2, and consider the latter possibility here. This would nonlocally create a particle in the unopened box B_2, and that action at a distance would have the following properties:

---

[13] This last example is given by Bell in his interview with Jeremy Bernstein [11, p. 63].

1. *That action is instantaneous:*[14] the particle has to be entirely in box B_2, at the same time as we open box B_1.

2a. *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.

2b. *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.

3. *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 wave function in the other half-box (coming from the cutting in two of the same box) but not in any other half-box.

4. *That action cannot be used to transmit messages*: if we open box B_1, we learn what the wave function 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. 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 fashion, at least according to ordinary quantum mechanics. Since we have no way, by acting on one box, to choose in which of the two boxes the particle will be, it is impossible to use that experiment to send messages.

The impossibility of sending messages is sometimes taken to mean that there is nothing nonlocal going on. But nonlocality refers here to causal interactions as described (in principle) by physical theories. Messages are far more anthropocentric than that, and require that humans be able to control these interactions in order to communicate. As remarked by Maudlin, the Big Bang and earthquakes cannot be used to send messages, but they have causal effects nevertheless [51, pp. 136–137].

Let us compare this kind of nonlocality with the nonlocality in Newtonian gravity. The latter also allows actions at a distance: since the gravitational force depends on the distribution of matter in the universe, changing that distribution, say by moving my body, instantaneously affects all other bodies in the universe. That action at a distance has properties 1 and 2a, but not the others. Of course, its effect decreases with the distance, because of the inverse square law, and it affects all bodies at a given distance equally (it is not individuated). On the other hand, it can in principle be used to transmit messages: if I decide to choose, at every minute, to wave my arm or not to wave it, then one can use that choice of movements to encode a sequence of zeros and ones and, assuming that the gravitational effect can be detected, one can therefore transmit a message instantaneously and arbitrarily far (but the further away one tries to transmit it, the harder the detection). Of course, all this refers to Newton's

---

[14] Of course, instantaneity is not a relativistic notion, so let us say, instantaneous in the reference frame where both boxes are at rest. We will not discuss relativity here, see [51, pp. 136–137] for a discussion of the conflict between relativity and quantum nonlocality.

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, as we shall see, this is a major difference with the situation in quantum mechanics).

It is well known that Newton did not like this aspect of his own theory.[15] In a letter to the theologian Richard Bentley, he wrote:

> [...] that one body may act upon another at a distance through a vacuum without the mediation of any thing else [...] is to me so great an absurdity that I believe no man who has in philosophical matters any competent faculty of thinking can ever fall into it.
>
> Isaac Newton [56, 53]

Not surprisingly, Einstein also firmly rejected the idea of action at a distance. In his discussions with Max Born (see Subsect. 7.2), he 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 [19, p. 164]

In the example of the boxes, this means that opening one half-box cannot possibly influence the physical situation in the other half-box.

The Dutch physicist Hendrik Casimir underlined 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 [22], quoted in [8]

If everything is connected with everything through nonlocal actions, then science becomes impossible, because, in order to test scientific theories, one always need to assume that one can isolate some systems or some variables. However, luckily, the nonlocality in quantum mechanics does not go so far as to make isolated systems impossible to realize in practice.

---

[15] Newton thought that gravitation was mediated by particles moving at a finite speed, so that the effect of gravitation could not be instantaneous. See [52] for more details.

However, because of the problems linked with nonlocality, post-Newtonian physics has tried to eliminate property 1, while classical electromagnetism or the general theory of relativity have kept only property 2a and the negation of 4. And, due to special relativity, the combination of 1 and the negation of 4 allows in principle the sending of messages into one's own past,[16] so that, if 1 holds, 4 must also hold.

One may ask whether quantum mechanics proves that there are physical effects displaying properties 1–4. 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.[17] In order to prove nonlocality in the sense introduced here, i.e., a phenomenon having properties 1–4 above, we have to turn to a more sophisticated situation.

5 A Simple Proof of Nonlocality

Let us start with an anthropomorphic thought experiment, but which is completely analogous to what happens in real experiments and could even, in principle, be realized in the anthropomorphic form presented here, as we will show at the end of this section.

Two people, A (for Alice) and B (for Bob) 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" (the reason why we introduce 3 questions will become clear below). This experiment is repeated many times, with A and B meeting together each time in the middle of the room, and the questions and answers vary apparently at random. When A and B are together in the room, they can decide to follow whatever "strategy" they want in order to answer the questions, but the statistics of their answers, although they look random, must nevertheless satisfy two properties. We impose these properties because they are analogous to what happens in real physical experiments, but translating the experiments into an anthropomorphic language may help us see how paradoxical our final conclusions are.

The *first property* is that, when the same question is asked at X and Y, one always gets the same answer. How can that be realized? One obvious possibility is that A and B agree upon which answers they will give before moving toward 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.

Another possibility is that, when A reaches door X, she calls B and tells him which question was asked and the answer she gave. Then, of course, B can just give the same answer as A 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 superluminal action at a distance be-

[16] At least, according to the usual understanding of relativity. See [51, 8] for a more detailed discussion.

[17] Indeed, this is what happens in the de Broglie-Bohm theory, see [14, 29, 30, 46, 20].
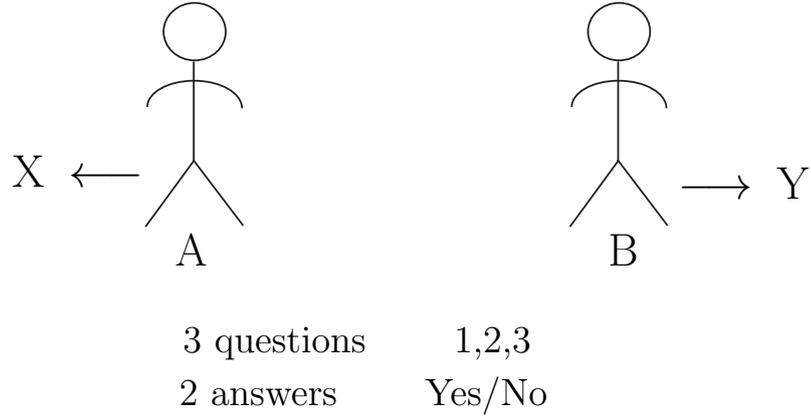
Fig. 3: The anthropomorphic experiment

tween A at X and B at Y. Maybe A and B communicate by telepathy! Of course, this is not to be taken seriously, but that is the sort of interaction that Einstein had in mind when he spoke of God using "telepathic methods" ([36] quoted in [27, p. 68]) or of "spooky actions at a distance" [19, 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 (assuming simultaneity of the answers) there is a "spooky action at a distance", namely a communication of some sort takes place between A and B *when* they are asked the questions. This is similar to the dilemma about the boxes: either the particle is in one of the boxes or there is some physical action between the two boxes.

Note that, to raise this dilemma, one question suffices instead of three: if the answers on both sides are always the same, then they must be predetermined if no communication is possible between the two sides. This dilemma can be called the EPR part of our argument.[18]

The reason we need three possible questions is that there is a *second property* of the statistics of the answers: when the two questions addressed to A and B are *different*, then the answers must be the same in only a quarter of the cases. And this property, combined with the idea that the properties are predetermined, leads to a contradiction. This is a very simple version of *Bell's theorem*[19]:

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

1 The answers are determined before the questions are asked and are the same on

_____

[18] However, one must mention that the original argument of Einstein, Podolsky, and Rosen was a bit different. Einstein, Podolsky, and Rosen considered two particles moving in opposite directions and such that, because of conversation laws, one could deduce the position or the velocity of one particle by measuring the position or the velocity of the other. This is similar to the answers to the questions discussed here, except for the fact that the answers here take two values (Yes/No) instead of the continuum of values taken by the possible positions and velocities. The version of the EPR argument used here is due to David Bohm [13].

[19] The argument given here is taken from [28]. In the original paper by Bell [3], the proof, although fundamentally the same, was more complicated.

both sides (for the same questions).

2 The frequency of having the same answers on both sides when the questions are different is 1/4.

Proof. 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
Y	Y	Y
Y	Y	N
Y	N	Y
Y	N	N
N	Y	Y
N	Y	N
N	N	Y
N	N	N

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

$$\text{Frequency (answer to } 1 = \text{answer to 2)}$$
$$+ \text{ Frequency (answer to } 2 = \text{answer to 3)}$$
$$+ \text{ Frequency (answer to } 3 = \text{answer to 1)} \geq 1 \,.$$

But if

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

we get $3/4 \geq 1$, which is a contradiction. ∎

The inequality saying that the sum of the frequencies is 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, and which is violated by the quantum predictions.

But before drawing conclusions from this theorem, let us briefly describe how the Alice and Bob could realize these "impossible" statistics. According to quantum mechanics, one can send two particles in opposite directions and with a wave function such that, if one measures the spin of those particles along the same direction on both sides (at X and Y), one always obtains the same answer.[20] And one can choose three

---

[20] To be precise, the spins are anti-correlated: if one measures the spin of the particle at X and one finds that it is up, the spin of the particle at Y will be down. But one can decide that an up spin at X means that the answer is yes on that side, and that an up spin at Y means that the answer is no on that side. In that way, the perfect anti-correlations are translated into getting always the same answers when one measures the spin along the same direction on both sides. One should also remark that real experiments are made with photons, where polarization plays the role played by the spin here.

different directions in such a way that the frequency of having the same answers on both sides when the directions along which the spin is measured are different is 1/4.

Then Alice and Bob can just arrange things so that such correlated particles are sent to them when they reach X and Y. Then, they measure the spin in directions corresponding to the questions being asked to them, and give their answers according to the results of their measurements.

It is easy to check that the nonlocality proven here does have the four properties listed in Sect. 4: the results are in principle obtained instantaneously (or at least the time between them is much less that the one taken by the light to go from X to Y), the correlations between the results obtained by Alice and Bob exist no matter how far Alice and Bob are from each other and do not depend on their distance; those correlations are between two people initially together (or two particles coming from a common source) and not between one object and the rest of the Universe (as would be the case for the gravitational force) and, since neither side can control what the results will be,[21] these strange correlations cannot be used to transmit messages.

6   Conclusions of the EPR-Bell argument

Taken by itself and forgetting about the EPR argument, Bell's result can be called a "no hidden variables theorem". Indeed, Bell showed that the mere supposition that the values of the spin pre-exist their "measurement" (remember that those values are called hidden variables, since they are not included in the usual quantum description), combined with the perfect anti-correlation when the axes along which measurements are made are the same, and the 1/4 result for the frequencies of correlations when measurements are made along different axes, leads to a contradiction. Since the last two claims are empirical predictions of quantum mechanics that have been amply verified (in a somewhat different form), this means that these hidden variables or pre-existing values cannot exist.[22]

But Bell, of course, always presented his result *in combination with* the EPR argument, which shows that the assumption of locality, combined with the perfect correlation when the directions of measurement (or questions) are the same, implies the existence of those hidden variables shown by Bell to be impossible, because their mere existence leads to a contradiction. 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.

To repeat, the EPR part of the argument shows that, if there are no pre-existing values, then the perfect correlations when the directions are the same imply some action at a distance. The Bell part of the argument, i.e., the theorem of the previous

---

[21] We will not give the proof of this fact, but is a well established consequence of the quantum formalism; see [31, 43, 5], [15, p. 139], or [51, Chap. 4].

[22] Which implies that, when physicists measure the spin of a particle, they do not measure any pre-existing property of that particle. This is nicely accounted for in the de Broglie-Bohm theory, see [14, 29, 30, 46, 20].

section, 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 directions are different.

But what does this mean? It means that some action at a distance does exist in Nature, but it does not tell us what this action consists of. And we cannot answer that question without having a theory that goes beyond ordinary quantum mechanics.[23]

In ordinary quantum mechanics, what is nonlocal is the "collapse of the wave function": when a measurement of the spin is made at X along a given direction, the wave function instantaneously collapses to a certain wave function depending on the result of the measurement but it does that *both at X and Y, so nonlocally.*

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 is no physical effect whatsoever or, if one interprets the collapse of the wave function as a mere gain of information, then this means that we must have those predetermined values that lead to a contradiction.

Given the radical nature of the conclusions of the EPR–Bell argument, many attempts have been made to avoid them, i.e., to claim that the world is local after all. One such 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 (see for example [65, 41]). In the same vein, one sometimes claims that science limits itself to predictions of empirical correlations, not to explanations.

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 [9, p. 152]

Another variant of the "shrugging one's shoulders" argument, is to invoke a sort of "conspiracy": for example, that each person has an answer to only one question but that, each time, and no matter how many times the experiment is repeated, that happens to be the question that is being asked to him or her. If we make that assumption,

---

[23] The de Broglie-Bohm theory is such a theory, and non-locality is to some extent explained in that theory, see [14, 29, 30, 46, 20].

then our theorem cannot be derived (for the proof of the theorem to work, we need to assume pre-existing answers for three questions).[24]

That move can be considered as an instance of what is known as the Duhem–Quine thesis in philosophy of science: no matter what the data are, one can always save one's favorite theory (here it would be 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". As the authors of [45] observe, "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" [45, 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 will always remain a logical possibility.

In any case, refusing to face a problem is not the same thing as solving it. One thing is certain: nobody has yet proposed a local explanation for those perfect correlations, and indeed nobody could do so, since Bell has shown that it is impossible.

Finally, one should emphasize that Einstein's speculations, which looked purely philosophical or even "metaphysical" to many, have led to what is probably "the most profound discovery of science", to use Henri Stapp's apt phrase [62, p. 271], namely the existence of nonlocal effects in the world. And the EPR and Bell papers laid the foundation for the quantum information revolution. This should be a lesson for "pragmatists".

On the other hand, it is ironical that this result refutes Einstein's most basic assumption about Nature. As Bell said, the question raised by EPR was answered in a way that Einstein would probably "have liked least", by showing that the "obvious" assumption of locality made by EPR is actually false [4, p. 11].

Let us now see how the physics community has, in general, misunderstood Einstein's (or EPR's) and Bell's arguments.

7   Misunderstandings of Einstein

One rather common misconception is to think that the goal of EPR was to show that one could measure, say, the spin in one direction at X and in another direction at Y for particles in the wave function described at the end of Sect.5, and therefore *know* the values of the spin of both particles in two different directions. Since the spin along two different axes cannot be measured simultaneously, according to ordinary quantum

---

[24] Another suggestion sometimes made is that the ordinary rules of probability do not apply in the EPR–Bell situation. But, as pointed out by Tumulka [64], since the reasoning here relies only on frequencies of results of experiments, and since the latter obviously do satisfy the ordinary rules of probability, this attempt to "save locality", by trying to deny the implications of the EPR–Bell argument, does not work.

mechanics, this would mean that the goal of EPR would be to refute that limitation in our ability to perform measurements.

But, as we explained in Sect.5, the point of the EPR paper was not to claim that one could *measure* quantities that are impossible to measure simultaneously according to quantum mechanics, but rather that, if one can learn something about a physical system by making a measurement on a distant system, then, barring actions at a distance, that "something" must already be there before the measurement is made on the distant system.

The dilemma for Einstein was: either the quantum mechanical description of reality is complete, in which case there is something nonlocal going on in Nature, or Nature is local but the quantum mechanical description of reality is incomplete, and hidden variables have to be added to it. In his 1949 *Reply to criticisms*, Einstein did actually pose the question in the form of a dilemma, implied by the EPR argument:

> [. . .] the paradox forces us to relinquish one of the following two assertions:
>
> (1) the description by means of the ψ-function is complete,
>
> (2) the real states of spatially separated objects are independent of each other.

> Albert Einstein [60, p. 682]

Let us now see how Bohr and Born responded to Einstein.

7.1   Bohr and Einstein

Bohr's viewpoint not something everybody agrees on, but, from what I understand, he insisted that quantum mechanics impose limitations to how much we can know about the microscopic world. In order to talk about that world, we need to use "classical" concepts, meaning concepts that deal with macroscopic objects or experimental devices, and since one cannot measure mutually incompatible properties in a single experiment, one has to resort to "complementary" views of the microscopic world: if we set up one type of experiment, we can talk about, for example, positions, and through another experiment, about momenta; or we could talk about spins in different directions, one for each experimental setup. In each setup, we can have well-defined answers, but is is pointless to try to combine these different "complementary" aspects of reality into a coherent mental image.

The root of the difference between Einstein and Bohr is that Einstein was arguing at the level of ontology, of *what there is*, insisting that a complete description of physical systems *must* go beyond the description given by ordinary quantum mechanics, if the world is local. Bohr, on the other hand, was answering systematically at the level of epistemology, of *what we can know*. Bohr was not answering Einstein, he was simply not listening to his objections. This was already the case at the 1927 Solvay Conference, when Bohr said that he did not "understand what precisely is the point which Einstein wants to [make]." ([2, p. 442]).

But the discussion between Bohr and Einstein centered mostly on the 1935 EPR paper.

The "onslaught" of EPR "came down upon us as a bolt from the blue", wrote Léon Rosenfeld, who was in Copenhagen at that time and who reported EPR's argument to Bohr. The latter then worked intensely "week after week" on a reply.[25]. There is a widespread belief among physicists that Bohr came up with an adequate answer to the EPR paper [16]. But, unlike Born, whose misunderstanding of Einstein (see Subsect. 7.2 below) is clearly stated, Bohr's answer is hard to understand,[26] despite the fact that, according to Rosenfeld, one of his favorite aphorisms was a line from the poet Friedrich von Schiller [67, p. 143]: "Only fullness leads to clarity, and truth dwells in the depths."

So, what was Bohr's reply? EPR had written that:

> If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity,

> Albert Einstein, Boris Podolsky, Nathan Rosen [33], reprinted in [67, pp. 138–141]

This again means that if, by doing something at A, we can learn something about the physical situation at B, then what we learn must exist before we learn it, since A and B can be far apart. Here of course, EPR assumed locality.

Bohr replied :

> [...] the wording of the above-mentioned criterion [...] contains an ambiguity as regards the meaning of the expression "without in any way disturbing a system". Of course there is in a case like that just considered no question of a mechanical disturbance of the system under investigation during the last critical stage of the measuring procedure. But even at this stage there is essentially the question of *an influence on the very conditions which define the possible types of predictions regarding the future behavior of the system* [...] their argumentation does not justify their conclusion that quantum mechanical description is essentially incomplete [...] This description may be characterized as a rational utilization of all possibilities of unambiguous interpretation of measurements, compatible with the finite and uncontrollable interaction between the objects and the measuring instruments in the field of quantum theory.

> Niels Bohr [16], quoted in [9, p. 155] (italics in the original)

---

[25] See Rosenfeld's commentary in [67, p. 142]

[26] As an amusing side remark, we note that, in the reprint of Bohr's paper in [67], which was the most accessible source of that paper before Internet, pages 148 and 149 are the wrong way round.

Bell dissects this passage as follows:

> Indeed I have very little idea what this means. I do not understand in what sense the word 'mechanical' is used, in characterizing the disturbances which Bohr does not contemplate, as distinct from those which he does. I do not know what the italicized passage means — 'an influence on the very conditions [...]'. Could it mean just that different experiments on the first system give different kinds of information about the second? But this was just one of the main points of EPR, who observed that one could learn *either* the position *or* the momentum of the second system.[27] And then I do not understand the final reference to 'uncontrollable interactions between measuring instruments and objects', it seems just to ignore the essential point of EPR that in the absence of action at a distance, only the first system could be supposed disturbed by the first measurement and yet definite predictions become possible for the second system. Is Bohr just rejecting the premise — 'no action at a distance' — rather than refuting the argument ?
>
> John Bell [9, pp. 155–156] (italics in the original)

Actually, Einstein also thought that Bohr rejected EPR's premise that "the real situation of B could not be influenced (directly) by any measurement taken on A" (see [60, pp. 681–682]). But this is by no means clearly stated and one may doubt that Bohr really understood Einstein.

Indeed, it is interesting to observe, as the physicist Howard Wiseman does:

> When reviewing the Einstein–Bohr debates in 1949 [17], Bohr concluded his summary of his reply to EPR by quoting his defence based upon complementarity. Astonishingly, he immediately followed this by an apology for his own 'inefficiency of expression which must have made it very difficult to appreciate the trend of the argumentation [...]'! But rather than taking the opportunity to explain himself more lucidly, he instead directed the reader to concentrate on the earlier debates with Einstein regarding the consistency of quantum mechanics.
>
> Howard Wiseman [69]

We may also note also that, in his 1949 "Reply to criticisms" [38], Einstein does not address Bohr's criticisms in any detail. Maybe this was due to the fact that he did not have a very high opinion of Bohr's "philosophy", at least that is the impression that one gets by reading his letters to Schrödinger, where one finds for example:

---

[27] Here Bell refers to the original EPR argument discussed in footnote 17 in Sect. 5. (Note by J.B.)

> The Heisenberg-Bohr tranquilizing philosophy – or religion? – is so delicately contrived that, for the time being, it provides a gentle pillow for the true believer from which he cannot very easily be aroused.
>
> <div align="right">Albert Einstein [32]</div>

In another letter to Schrödinger, Einstein referred to Bohr as the "Talmudic philosopher" for whom "reality is a frightening creature of the naive mind" [34]. Einstein also referred to Bohr as "the mystic, who forbids, as being unscientific, an enquiry about something that exists independently of whether or not it is observed,...". [35].

7.2 Born and Einstein

Born and Einstein were lifelong friends; they both had to leave Germany, to avoid persecution, Born settling in Edinburgh, Einstein in Princeton. They exchanged a long correspondence, about physics, personal matters, and politics, which was edited by Born in 1971, long after Einstein's death in 1955 [19]. In that correspondence, they also discussed quantum mechanics. It is rather fascinating to read their exchange of letters and to observe the systematic misunderstanding of Einstein's point of view by Born.

For example, Einstein wrote an article in 1948 and sent it to Born (it is reproduced in the correspondence), begging him to read it as if he had just arrived "as a visitor from Mars". Einstein hoped that the article would help Born to "understand my principal motives far better than anything else of mine you know" [19, p. 168]. In that article, Einstein restated his unchanging criticisms:

> If one asks what, irrespective of quantum mechanics, is characteristic of the world of ideas of physics, one is first of all struck by the following: the concepts of physics relate to a real outside world, that is, ideas are established relating to things such as bodies, fields, etc., which claim a 'real existence' that is independent of the perceiving subject — ideas which, on the other hand, have been brought into as secure a relationship as possible with the sense-data. It is further characteristic of these physical objects that they are thought of as arranged in a spacetime continuum. An essential aspect of this arrangement of things in physics is that they lay claim, at a certain time, to an existence independent of one another, provided these objects 'are situated in different parts of space'. [...]
>
> The following idea characterizes the relative independence of objects far apart in space (A and B): external influence on A has no direct influence on B.
>
> <div align="right">Albert Einstein [37], reproduced in [19, pp. 170–171]</div>

Then Einstein repeats his usual argument, discussed above, that different choices of measurement at A create (through the collapse rule) a different wave function at B.

Here is Born's reaction to that very same article:

> The root of the difference between Einstein and me was the axiom that events which happens in different places A and B are independent of one another, in the sense that an observation on the states of affairs at B cannot teach us anything about the state of affairs at A.

> Max Born [19, p. 176]

As Bell says:

> Misunderstanding could hardly be more complete. Einstein had no difficulty accepting that affairs in different places could be correlated. What he could not accept was that an intervention at one place could *influence*, immediately, affairs at the other.

> These references to Born are not meant to diminish one of the towering figures of modern physics. They are meant to illustrate the difficulty of putting aside preconceptions and listening to what is actually being said. They are meant to encourage *you*, dear listener, to listen a little harder.

> John Bell [9, p. 144] (italics in the original)

What Born said was that making an experiment in one place *teaches* us something about what is happening at another place, which is unsurprising. If, in the anthropomorphic example given in Sect. 5, both people had agreed on a common strategy, one would learn what B would answer to question 1, 2, or 3, by asking A that same question. But, and that was Einstein's point, it would mean that the answers were predetermined and, thus, that quantum mechanics was incomplete.

So, it seems that Born did in fact agree with Einstein that quantum mechanics is incomplete, but simply did not understand what Einstein meant by that.

8 Misunderstandings of Bell

As we saw, if one forgets about the EPR argument, Bell's result, can be stated as a "no hidden variables theorem". But for Bell, his result, combined with the EPR result was not a "no hidden variables theorem", but a nonlocality theorem, the result on the impossibility of hidden variables being only one step in a two-step argument.

Bell was perfectly clear about this:

> Let me summarize once again the logic that leads to the impasse. The EPRB correlations[28] are such that the result of the experiment on one side immediately foretells that on the other, whenever the analyzers happen to be parallel. If we do not accept the intervention on one side as a causal

---

[28] Here EPRB means EPR and Bohm, who reformulated the EPR argument in term of spins [13], which is the formulation used here and in most of the literature. (Note by J.B.)

influence on the other, we seem obliged to admit that the results on both sides are determined in advance anyway, independently of the intervention on the other side, by signals from the source and by the local magnet setting. But this has implications for non-parallel settings which conflict with those of quantum mechanics. So we *cannot* dismiss intervention on one side as a causal influence on the other.

<div align="right">John Bell [9, pp. 149–150] (italics in the original)</div>

He was also conscious of the misunderstandings of his results:

It is important to note that to the limited degree to which *determinism*[29] plays a role in the EPR argument, it is not assumed but *inferred*. What is held sacred is the principle of "local causality" — or "no action at a distance". [...]

It is remarkably difficult to get this point across, that determinism is not a *presupposition* of the analysis.

<div align="right">John Bell [9, p. 143] (italics in the original)</div>

And he added, unfortunately only in a footnote:

My own first paper on this subject (*Physics* **1**, 195 (1965))[30] starts with a summary of the EPR argument *from locality to* deterministic hidden variables. But the commentators have almost universally reported that it begins with deterministic hidden variables.

<div align="right">John Bell [9, p. 157, footnote 10] (italics in the original)</div>

An example of such a commentator is the famous physicist Murray Gell-Mann, Nobel Prize winner and discoverer of the quarks, who wrote:

Some theoretical work of John Bell revealed that the EPRB experimental setup could be used to distinguish quantum mechanics from hypothetical hidden variable theories [...] After the publication of Bell's work, various teams of experimental physicists carried out the EPRB experiment. The result was eagerly awaited, although virtually all physicists were betting on the corrections of quantum mechanics, which was, in fact, vindicated by the outcome.

<div align="right">Murray Gell-Mann [42, p. 172]</div>

---

[29] Here, "determinism" refers to the idea of pre-existing values. (Note by J.B.)
[30] Reference [3], reprinted as Chap. 2 in [9]. (Note by J.B.)

So Gell-Mann opposes hidden variable theories to quantum mechanics, but the only hidden variables that Bell considered were precisely those that were needed, because of the EPR argument, in order to "save" locality. So if there is a contradiction between the existence of those hidden variables and experiments, it is not just quantum mechanics that is vindicated, but locality that is refuted.

In one of his most famous papers, "Bertlmann's socks and the nature of reality" [6], Bell tried to give a pedagogical argument to explain his main idea. He gave the example of a person (Reinhold Bertlmann) who always wears socks of different colors.[31] If we see that one sock is pink, we know automatically that the other sock is not pink (let's say it is green). That would be true even if the socks were arbitrarily far away. So by looking at one sock, we learn something about the other sock and there is nothing surprising about that, because socks *do have* a color whether we look at them or not.[32] But what would we say if we were told that the socks have no color before we look at them? That would be surprising, of course, but the idea that quantum mechanics is complete means exactly that (if we replace the color of the socks by the values of the spin before measurement). But then, looking at one sock would "create" not only the color of that sock but also the color of the other sock. And that implies the existence of actions at a distance, if the socks are far apart. Bell emphasized that in quantum mechanics, we are in a situation similar to this nonlocal creation of colors for the socks.

However, Murray Gell-Mann made the following comment on the Bertlmann's socks paper:

> The situation is like that of Bertlmann's socks, described by John Bell in one of his papers. Bertlmann is a mathematician who always wears one pink and one green sock. If you see just one of his feet and spot a green sock, you know immediately that his other foot sports a pink sock. Yet no signal is propagated from one foot to the other. Likewise no signal passes from one photon to the other in the experiment that confirms quantum mechanics. No action at a distance takes place. (The label "nonlocal" applied by some physicists to quantum-mechanical phenomena like the EPRB effect is thus an abuse of language. What they mean is that *if interpreted classically in terms of hidden variables*, the result would indicate nonlocality, but of course such a classical interpretation is wrong.)

> Murray Gell-Mann [42, pp. 172–173]

This is not correct: it is true that, because of the random nature of the results, the experimental setup of EPRB cannot be used to send messages (or signals), as we saw in Sect. 4. But nevertheless, some action at a distance does take place. And the remark

---

[31] See [9, p. 139] for a picture of Mr Bertlmann wearing one sock indicated "pink" and the other "not pink".

[32] In case some people might worry that the notion of "color" refers to our sensations, let us be more precise and say that the socks have the physico-chemical properties that will reflect a light of a certain wavelength (and the fact that they have those properties is independent of our perception), which will ultimately produce our perception of color.

about "classical interpretation" completely misses EPR's argument showing that hidden variables are necessary if locality holds.

Of course, Gell-Mann's goal in the passage quoted here is probably to dismiss pseudo-scientific exploitations of Bell's result (such as invoking it to justify the existence of telepathy), but his defense of science is misdirected: the behavior of quantum particles is *not* like that of Bertlmann's socks (which is indeed totally unsurprising), and that was the whole point of Bell's paper.

Eugene Wigner, another Nobel Prize winner, also saw Bell's result solely as a no hidden variables result: "In my opinion, the most convincing argument against the theory of hidden variables was presented by J.S. Bell." [68, p. 291].

This is misleading, because Wigner considers only Bell's argument, which indeed shows that pre-existing spin values (or "hidden variables") cannot exist, but forgets the EPR part of the argument, which was Bell's starting point.[33]

To give yet another example of misunderstanding of Bell, coming from a well-known defender of the Copenhagen interpretation, consider what Rudolf Peierls declared in an interview, referring to Aspect's experiments [1] that verified the violation of Bell's inequalities:

> If people are obstinate in opposing the accepted view they can think of many new possibilities, but there is no sensible view of hidden variables which doesn't conflict with these experimental results. That was proved by John Bell, who has great merit in establishing this. Prior to that there was a proof due to the mathematician von Neumann,[34] but he made an assumption which is not really necessary.

> Rudolf Peierls in [24, p. 77]

Basically the same mistakes were made in 1999 by one of the most famous physicists of our time, Stephen Hawking:

> Einstein's view was what would now be called a hidden variables theory. Hidden variables theories might seem to be the most obvious way to incorporate the Uncertainty Principle into physics.[35] They form the basis of the mental picture of the universe, held by many scientists, and almost all philosophers of science. But these hidden variable theories are wrong. The British physicist, John Bell, who died recently, devised an experimental test that would distinguish hidden variable theories. When the experiment was carried out carefully, the results were inconsistent with hidden variables.

> Stephen Hawking [48]

---

[33] See Goldstein [44] for a further discussion of Wigner's views.

[34] See Sect. 9 for a brief explanation of von Neumann's argument. (Note by J.B.)

[35] By this, Hawking, means that hidden variables would be a way to maintain determinism, despite the fact that precise measurements of both position and velocities are prohibited by Heisenberg's uncertainty principle. (Note by J.B.)

The misconceptions about Bell continue: in an article about a Symposium held in 2000 in Vienna, in honor of the 10th anniversary of John Bell's death, the physicist Gerhard Grössing wrote:

> In the 11 August 2000 issue of "Science", D. Kleppner and R. Jackiw of MIT published a review article on "One Hundred Years of Quantum Physics" [50]. They discuss Bell's inequalities and also briefly mention the possibility of "hidden variables". However, according to the authors, the corresponding experiments had shown the following: "Their collective data came down decisively against the possibility of hidden variables. For most scientists this resolved any doubt about the validity of quantum mechanics."
>
> Gerhard Grössing [47]

In the February 2001 issue of *Scientific American*, Max Tegmark and John Archibald Wheeler published an article entitled *100 Years of the Quantum* (quoted by Grössing in [47]), where one reads:

> Could the apparent quantum randomness be replaced by some kind of unknown quantity carried out inside particles, so-called 'hidden variables'? CERN theorist John Bell showed that in this case, quantities that could be measured in certain difficult experiments would inevitably disagree with standard quantum predictions. After many years, technology allowed researchers to conduct these experiments and eliminate hidden variables as a possibility.
>
> Max Tegmark and John Archibald Wheeler [63, p. 76]

These two last quotes clearly commit the usual mistake: they ignore the fact that Bell's result, combined with the EPR argument, refutes locality not merely the existence of "hidden variables".

In the preface to the proceedings of the Symposium held in honor of the 10th anniversary of John Bell's death, Reinhold Bertlmann (mentioned in the "socks" paper [6]) and the distinguished Austrian experimentalist Anton Zeilinger expressed succinctly what is probably the view of the majority of physicists about Bell's result, and which repeats the same mistake as in the previous quotes:

> [...] while John Bell had flung open the door widely for hidden variables theories [by showing the flaws in von Neumann's proof], he immediately dealt them a major blow. In 1964, in his celebrated paper "On the Einstein–Podolsky–Rosen Paradox", he showed that any hidden variables theory, which obeys Einstein's requirement of locality, i.e., no influence travelling faster than the speed of light, would automatically be in conflict with quantum mechanics. [...] While a very tiny [experimental] loophole in principle remains for local realism, it is a very safe position to assume that quantum

mechanics has definitely been shown to be the right theory. Thus, a very deep philosophical question, namely, whether or not events observed in the quantum world can be described by an underlying deterministic theory, has been answered by experiment, thanks to the momentous achievement of John Bell.

Reinhold Bertlmann and Anton Zeilinger, preface to [12]

David Mermin summarized the situation described in this section in an amusing way:

Contemporary physicists come in two varieties.

Type 1 physicists are bothered by EPR and Bell's theorem.

Type 2 (the majority) are not, but one has to distinguish two subvarieties.

Type 2a physicists explain why they are not bothered. Their explanations tend either to miss the point entirely (like Born's to Einstein)[36] or to contain physical assertions that can be shown to be false.

Type 2b are not bothered and refuse to explain why. Their position is unassailable. (There is a variant of type 2b who say that Bohr straightened out the whole business, but refuse to explain how.)

David Mermin [54]

However, the same David Mermin also wrote:

Bell's theorem establishes that the value assigned to an observable must depend on the complete experimental arrangement under which it is measured, even when two arrangements differ only far from the region in which the value is ascertained — a fact that Bohm theory[37] exemplifies, and that is now understood to be an unavoidable feature of any hidden-variables theory.

To those for whom nonlocality is anathema, Bell's Theorem finally spells the death of the hidden-variables program.

David Mermin [55, p. 814]

But Bell's theorem, combined with EPR, shows that nonlocality, whether we consider it anathema or not, is an unavoidable feature of the world.

---

[36] See Subsect. 7.2. (Note by J.B.)

[37] See [14, 29, 30, 46, 20] for an exposition of that theory, which is called in [20] the de Broglie–Bohm theory. (Note by J.B.)

Mermin adds in a footnote: "Many people contend that Bell's theorem demonstrates nonlocality independently of a hidden-variables program, but there is no general agreement about this." [55] It is true that there is no agreement (to put it mildly), but Mermin's own remark on the two varieties of physicists seems to explain why this is so.[38]

To conclude, let us highlight some more positive reactions to Bell's result. As we mentioned at the end of Sect. 6, Henry Stapp, a particle theorist at Berkeley, wrote that "Bell's theorem is the most profound discovery of science"[62, p. 271]. David Mermin mentions an unnamed "distinguished Princeton physicist" who told him [54]: "Anybody who's not bothered by Bell's theorem has to have rocks in his head." And Feynman derived in 1982 an inequality similar to Bell's (no reference to Bell's work is made in his paper), from which he drew the following conclusion:

> That's all. That's the difficulty. That's why quantum mechanics can't seem to be imitable by a local classical computer.[39] I've entertained myself always by squeezing the difficulty of quantum mechanics into a smaller and smaller place, so as to get more and more worried about this particular item. It seems to be almost ridiculous that you can squeeze it to a numerical question that one thing is bigger than another.
>
> Richard Feynman [39, p. 485]

What Feyman calls "one thing being bigger than another" are inequalities of the type $1 \geq 3/4$ that we used in Sect. 5.

9 Conclusions

We have only discussed parts of the misunderstandings that are rather widespread both among physicists and in the lay public about the history of quantum mechanics. For example, Schrödinger thought that his famous cat was a *reductio ad absurdum* of the idea that quantum mechanics was complete: for him, the cat was alive or dead, not both, and since the pure quantum mechanical description of the cat does not specify which is the case (the cat is in a "superposed state"), it is manifestly incomplete or, in other words, the cat being alive or dead is another example of introducing "hidden variables" (variables that are not part of the wave function), although in this case they are not hidden at all! Now, many people seem to think that quantum mechanics has shown that the cat is both alive and dead before anybody looks at the situation, which is the exact opposite of what Schrödinger tried to show.

Another example of misunderstandings is provided by von Neumann's no hidden variable theorem. In his 1932 book [66], von Neumann tried to show that no hidden variables could be introduced that would complete the quantum mechanical description.

---

[38] What is even more surprising is that this comment comes at the end of a remarkably clear paper on the no hidden variables theorems [55].

[39] Which means what we call locality here. (Note by J.B.)

Although his proof was technically correct (and rather simple), it relied on assumptions that were quite arbitrary from a physical point of view, as shown later by Bell [7]. But, because of the enormous (and justified) prestige of von Neumann as a mathematician, most physicists (including for example Max Born in [18, p. 109] or in the quote of Peierls cited in Sect. 8) did not bother to look closely at his assumptions. Nowadays, von Neumann's theorem against the possibility of hidden variables is less often cited, but for a long time it was considered as the last nail in the coffin of the mere idea that such variables might be introduced.

The lesson of this sad history is that we are unlikely to understand the real issues raised by quantum mechanics, whether it be the role of the observer or nonlocality, if we continue to adhere to a history of quantum mechanics written by the victors (at least until now) and that systematically misrepresents the views of Einstein, Schrödinger, de Broglie, Bohm and Bell. Reading those authors without prejudice would already constitute a big step in the right direction.

References

[1] A. Aspect, J. Dalibard and G. Roger: Experimental test of Bell's inequalities using time-varying analysers, *Physical Review Letters* **49**, 1804–1807 (1982)

[2] G. Bacciagaluppi and A. Valentini: *Quantum Mechanics at the Crossroads. Reconsidering the 1927 Solvay Conference*, Cambridge University Press, Cambridge, 2009

[3] J.S. Bell: On the Einstein–Podolsky–Rosen paradox, *Physics* **1**, 195–200 (1964). Reprinted as Chap. 2 in [9]

[4] J.S. Bell: On the problem of hidden variables in quantum mechanics, *Reviews of Modern Physics* **38**, 447–452 (1966). Reprinted as Chap. 1 in [9]

[5] J.S. Bell: The theory of local beables, *Epistemological Letters*, March 1976. Reprinted as Chap. 7 in [9]

[6] J.S. Bell: Bertlmann's socks and the nature of reality, *Journal de Physique* **42** C2, 41–61 (1981). Reprinted as Chap. 16 in [9]

[7] J.S. Bell: On the impossible pilot wave, *Foundations of Physics* **12**, 989–999 (1982). Reprinted as Chap. 17 in [9]

[8] J.S. Bell: La nouvelle cuisine. In: *Between Science and Technology*, A. Sarlemijn and P. Kroes (eds.), Elsevier Science Publishers, Amsterdam, 1990. Reprinted as Chap. 24 in [9]

[9] J.S. Bell: *Speakable and Unspeakable in Quantum Mechanics. Collected Papers on Quantum Philosophy*, 2nd edn, with an introduction by Alain Aspect, Cambridge University Press, Cambridge, 2004; 1st edn 1993

[10] M. Bell and S. Gao (eds.): *Quantum Nonlocality and Reality. 50 Years of Bell's Theorem*, Cambridge University Press, Cambridge, 2016

[11] J. Bernstein: *Quantum Profiles*, Princeton University Press, Princeton, 1991

[12] R.A. Bertlmann and A. Zeilinger (eds): *Quantum [Un]speakables From Bell to Quantum Information*, Springer, Berlin, 2002

[13] D. Bohm: *Quantum Theory*, Dover Publications, New York, New edition, 1989; First edition: Prentice Hall, Englewood Cliffs (N.J.), 1951

[14] D. Bohm: A suggested interpretation of the quantum theory in terms of "hidden variables", Parts 1 and 2, *Physical Review* **89**, 166–193 (1952). Reprinted in [67] pp. 369–390

[15] D. Bohm and B.J. Hiley: *The Undivided Universe*, Routledge, London, 1993

[16] N. Bohr: Can quantum mechanical description of reality be considered complete? *Physical Review* **48**, 696–702 (1935)

[17] N. Bohr: Discussion with Einstein on epistemological problems in atomic physics. In: [60], pp. 201–241

[18] M. Born: *Natural Philosophy of Cause and Chance*, Clarendon, Oxford, 1949

[19] M. Born (ed.): *The Born–Einstein Letters*, Macmillan, London, 1971

[20] J. Bricmont, *Making Sense of Quantum Mechanics*, Springer, Berlin, 2016

[21] J. Bricmont, What Did Bell Really Prove? In: [10, pp. 49-78]

[22] H.B.G. Casimir: In: Koninklijke Nederlandse Akademie van Wetenschappen 1808–1958, Noord-Hollandsche Uitgeversmij., Amsterdam 1958, pp. 243–51

[23] J.T. Cushing and E. MacMullin (eds): *Philosophical Consequences of Quantum Theory. Reflections on Bell's Theorem*, University of Notre Dame Press, Notre Dame, 1989

[24] P.C.W. Davies and J.R. Brown (eds): *The Ghost in the Atom: A Discussion of the Mysteries of Quantum Physics*, Cambridge University Press, Cambridge, 1993

[25] L. de Broglie: L'interprétation de la mécanique ondulatoire, *Le Journal de Physique et le Radium* **20**, 963–979 (1959)

[26] L. de Broglie: *The Current Interpretation of Wave Mechanics: A Critical Study*, Elsevier, Amsterdam, 1964

[27] H. Dukas, B. Hoffmann (eds): *Albert Einstein: The Human Side*, Princeton University Press, Princeton, 1979

[28] D. Dürr, S. Goldstein, R. Tumulka and N. Zanghì: John Bell and Bell's theorem. In: *Encyclopedia of Philosophy*, D.M. Borchert ed., Macmillan Reference, USA, 2005

[29] D. Dürr, S. Teufel, *Bohmian Mechanics. The Physics and Mathematics of Quantum Theory*, Springer, Berlin-Heidelberg, 2009.

[30] D. Dürr, S. Goldstein and N. Zanghì: *Quantum Physics Without Quantum Philosophy*, Springer, Berlin, Heidelberg, 2012

[31] P.H. Eberhard: Bell's theorem and the different concepts of locality, *Il Nuovo Cimento* **46B**, 392–419 (1978)

[32] A. Einstein: Letter to Erwin Schrödinger, 31 May 1928, in [59, p. 31]

[33] A. Einstein, B. Podolsky and N. Rosen: Can quantum mechanical description of reality be considered complete?, *Physical Review* **47**, 777–780 (1935)

[34] A. Einstein: Letter to Erwin Schrödinger, 19 June 1935, in [40, p. 35]

[35] A. Einstein: Letter to Erwin Schrödinger, 9 August 1939, in [59, p. 36]

[36] A. Einstein: Letter to Cornel Lanczos, 21 March 1942, in [27, p. 68]

[37] A. Einstein: Quantum mechanics and reality, *Dialectica* **2**, 320–324 (1948)

[38] A. Einstein: Remarks concerning the essays brought together in this co-operative volume. In [60, pp. 665–688]

[39] R. Feynman, Simulating physics with computers, *International Journal of Theor. Phys.*, **21**, 467-488, 1982.

[40] A. Fine: *The Shaky Game: Einstein Realism and the Quantum Theory*, University of Chicago Press, Chicago, 1986

[41] A. Fine: Do correlations need to be explained? In [23, pp. 175–194] (1989)

[42] M. Gell-Mann: *The Quark and the Jaguar*, Little, Brown and Co., London, 1994

[43] G.C. Ghirardi, A. Rimini and T. Weber: A general argument against superluminal transmission through the quantum mechanical measurement process, *Lettere Al Nuovo Cimento* **27**, 293–298 (1980)

[44] S. Goldstein: Quantum philosophy: the flight from reason in science. In: *The Flight from Science and Reason*, P. Gross, N. Levitt, and M.W. Lewis (eds), *Annals of the New York Academy of Sciences* **775**, 119–125 (1996). Reprinted in [30, Chap. 4]

[45] S. Goldstein, T. Norsen, D.V. Tausk and N. Zanghì: Bell's theorem, *Scholarpedia* 6(10): 8378 (2011)

[46] S. Goldstein: Bohmian mechanics, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.) (Spring 2013 Edition) `plato.stanford.edu/archives/spr2013/entries/qm-bohm/`

[47] G. Grössing: Serious Matter: The John Bell Scandal, available from `www.nonlinearstudies.at/Bell_E.php`

[48] S. Hawking: Does God play dice? (1999). Available online: `www.hawking.org.uk/doesgod-play-dice.html`

[49] W. Heisenberg: *The Physical Principles of Quantum Theory*, translated by C. Eckart and F. Hoyt, Dover publications, New York, 1949. First edition: University of Chicago Press, Chicago, 1930

[50] D. Kleppner and R. Jackiw: One hundred years of quantum physics, *Science* **289**, 893–898 (2000)

[51] T. Maudlin: *Quantum Nonlocality and Relativity*, Blackwell, Cambridge, 1st edn, 1994, 3rd edn, 2011

[52] T. Maudlin: What Bell did, *Journal of Physics A: Math. Theor.* **47**, 424010 (2014)

[53] E. McMullin: The explanation of distant action: Historical notes. In [23, pp. 272–302]

[54] D. Mermin: Is the moon there when nobody looks? Reality and the quantum theory, *Physics Today* **38**, April 1985, 38–47

[55] D. Mermin: Hidden variables and the two theorems of John Bell, *Reviews of Modern Physics* **65**, 803–815 (1993)

[56] I. Newton: Letter to Richard Bentley, 25 February 1693. In: *The Correspondence of Isaac Newton*, Vol. 3, H.W. Turnbull (ed.), Cambridge University Press, Cambridge, 1959, pp. 253–254

[57] T. Norsen: Einstein's boxes, *American Journal of Physics* **73**, 164–176 (2005)

[58] A. Pais: Einstein and the quantum theory, *Reviews of Modern Physics* **51**, 863–914, 1979

[59] K. Przibram (ed.): *Letters on Wave Mechanics, Schrödinger, Planck, Einstein; Lorentz.* Translated and with an Introduction by M.J. Klein, Philosophical Library, New York, 1967

[60] P.A. Schilpp (ed.): *Albert Einstein, Philosopher-Scientist*, The Library of Living Philosophers, Evanston, Illinois, 1949

[61] *Electrons et photons: rapports et discussions du 5e Conseil de Physique tenu à Bruxelles du 24 au 29 octobre 1927*, Instituts Solvay, Gauthier-Villars et cie, Paris, 1928. English translation in [2]

[62] H.P. Stapp: Bell's theorem and world process, *Il Nuovo Cimento* **29B**, 270 (1975)

[63] M. Tegmark and J.A. Wheeler: 100 years of quantum mysteries, *Scientific American*, February 2001, 72–79

[64] R. Tumulka: The assumptions of Bell's proof. In: [10, pp. 79-90]; and in: *International Journal of Quantum Foundations* (electronic journal). Available at `www.ijqf.org/groups-2/bells-theorem/forum/topic/the-assumptions-of-bells-proof/` (2014)

[65] B.C. van Fraasen: The charybdis of realism: Epistemological implications of Bell's inequality, *Synthese* **52**, 25–38 (1982). Reprinted in [23, pp. 97–113]

[66] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton, 1955. First edition in German, *Mathematische Grundlagen der Quantenmechanik*, 1932

[67] J.A. Wheeler and W.H. Zurek (eds): *Quantum Theory and Measurement*, Princeton University Press, Princeton, 1983

[68] E.P. Wigner: Interpretation of quantum mechanics. In: [67, pp. 260–314]

[69] H.M. Wiseman: From Einstein's theorem to Bell's theorem: A history of quantum nonlocality, *Contemporary Physics* **47**, 79–88 (2006)