John Bell and Bell's Theorem

Detlef Dürr, Sheldon Goldstein,
Roderich Tumulka, and Nino Zanghì

December 27, 2004

John Stewart Bell (1928-1990), a truly deep and serious thinker, was one of the leading physicists of the 20th century. He became famous for his discovery that quantum mechanics implies that nature is nonlocal, i.e., that there are physical influences between events that propagate faster than light.

From 1960 until his death Bell worked at CERN in Geneva on the physics of particle accelerators, making a number of important contributions to high-energy physics and quantum field theory. Noteworthy was his discovery in 1969, together with Roman Jackiw, of the so-called "Bell-Jackiw-Adler" anomaly (discovered independently by Stephen Adler), a mechanism explaining physical effects such as neutral pion decay (which are unexplainable on the basis of the symmetries of the classical field Lagrangian), in terms of an "anomalous" term arising from the renormalization of quantum field theory. Since then this mechanism has become an important cornerstone of quantum field theory. Another important contribution was the argument he gave in 1967 for why weak interactions should be described using a gauge theory.

John Bell was one of the leading experts—perhaps *the* leading expert—on the foundations of quantum mechanics. The book collecting his articles on this subject, *Speakable and unspeakable in quantum mechanics*, is unsurpassed for clarity and depth and it is still the best reference for whoever wishes to learn about the field [4].

Bell strongly opposed the "Copenhagen interpretation" of quantum physics, according to which macroscopic objects, such as chairs and planets, do exist out there, but electrons and other microscopic particles do not. According to the Copenhagen view, the world is divided into two realms, macro and micro, "classical" and "quantum," logical and contradictory—or, as Bell put it in one of his essays, into "speakable" and "unspeakable." Along with Albert Einstein, Erwin Schrödinger, Louis de Broglie and →David Bohm, Bell was one of the few physicists compelled by his conscience to reject the Copenhagen interpretation.

Bell emphasized that the empirical facts of quantum physics do not at all force us to renounce realism. There is, in fact, a realist theory (→Bohmian mechanics, also known as the de Broglie-Bohm theory) that accounts—insofar as the nonrelativistic theory is concerned—for all of these facts in a most elegant way. This theory describes a world in which electrons, quarks and the like are point particles, always having positions that move in a manner dictated by the wave function. It should be taught to students, Bell insisted, as a legitimate alternative to the prevailing orthodoxy. After GianCarlo Ghirardi, Alberto Rimini, and Tullio Weber in 1986 succeeded in formulating a second kind of realist theory, Bell encouraged the further development of this theory as well [4, p. 201]. He thought that such a theory contained the seeds of a reconciliation of quantum mechanics with fundamental Lorentz invariance, and thus a resolution of the tension between quantum mechanics and relativity that arose from his own work on quantum nonlocality.

In 1964, Bell proved that any serious version of quantum theory (regardless of whether or not it is based on microscopic realism) must violate locality [2]. He showed that if nature is governed by the predictions of quantum theory, the "locality principle," precluding any sort of instantaneous (or superluminal) action-at-a-distance, is simply wrong, and our world is nonlocal. The theoretical analysis leading to such a conclusion is commonly known as *Bell's theorem*.

Bell's theorem involves two parts. The first part is the Einstein-Podolsky-Rosen argument [6] applied to the simplified version considered by David Bohm [5], the EPRB experiment: a pair of spin one-half particles, prepared in a spin-singlet state, are moving freely in opposite directions. Measurements are made, say by Stern-Gerlach magnets, on selected components of the spins of the two particles. The spin-singlet state has the following property: whenever the component of the spin $\boldsymbol{\sigma}_1$ in any direction $\boldsymbol{\alpha}$ is measured for one of the two particles, a measurement of the same component of the spin $\boldsymbol{\sigma}_2$ of the other particle will give with certainty the opposite value. For such a state the assumption of locality implies the existence of what are often called noncontextual hidden variables. More precisely, it implies, for the spin-singlet state, the existence of random variables $Z^i_{\boldsymbol{\alpha}}$ $(= Z_{\boldsymbol{\alpha}\cdot\boldsymbol{\sigma}_i})$, $i = 1, 2$, which can be regarded as corresponding to preexisting values of all possible spin components of the two particles. In particular, focusing on components in only 3 directions $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ for each particle, locality implies the existence of 6 random variables $Z^i_{\boldsymbol{\alpha}}$, $i = 1, 2$, $\boldsymbol{\alpha} = \mathbf{a},\ \mathbf{b},\ \mathbf{c}$ such that

$$Z^i_{\boldsymbol{\alpha}} = \pm 1 \tag{1}$$

$$Z^1_{\boldsymbol{\alpha}} = -Z^2_{\boldsymbol{\alpha}} \tag{2}$$

and, more generally,

$$\mathrm{Prob}(Z^1_{\boldsymbol{\alpha}} \neq Z^2_{\boldsymbol{\beta}}) = q_{\boldsymbol{\alpha}\boldsymbol{\beta}}, \tag{3}$$

where the $q_{\boldsymbol{\alpha}\boldsymbol{\beta}} = (1 + \boldsymbol{\alpha}\cdot\boldsymbol{\beta})/2 = \cos^2(\theta/2)$ are the corresponding quantum mechanical probabilities, with θ the angle between $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$.

The argument for this conclusion can be expressed as follows: The existence of such random variables amounts to the idea that measurements of the spin components reveal preexisting values (the $Z^i_{\boldsymbol{\alpha}}$). Assuming locality, this is implied by the perfect quantum mechanical anticorrelations [2]:

> Now we make the hypothesis, and it seems one at least worth considering, that if the two measurements are made at places remote from one another the orientation of one magnet does not influence the result obtained with the other. Since we can predict in advance the result of measuring any chosen component of $\boldsymbol{\sigma}_2$, by previously measuring the same component of $\boldsymbol{\sigma}_1$, it follows that the result of any such measurement must actually be predetermined.

Otherwise the result would have, at least in part, been produced by the remote measurement, just the sort of influence that Bell's locality hypothesis precludes. We may also note that if the results had not been predetermined, the widely separated correlated residual innovations thereby implied would be an instance of nonlocality.

Observe that, given locality, the existence of such variables is a consequence rather than an assumption of Bell's analysis. Bell repeatedly stressed this point (by determinism Bell here means the existence of the preexisting values that would determine the results of the corresponding measurements):

> It is important to note that to the limited degree to which *determinism* 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. ([4], p. 143)
>
> Despite my insistence that the determinism was inferred rather than assumed, you might still suspect somehow that it is a preoccupation with determinism that creates the problem. Note well then that the following argument makes no mention whatever of determinism. ... Finally you might suspect that the very notion of particle, and particle orbit ... has somehow led us astray. ... So the following argument will not mention particles, nor indeed fields, nor any other particular picture of what goes on at the microscopic level. Nor will it involve any use of the words 'quantum mechanical system', which can have an unfortunate effect on the discussion. The difficulty is not created by any such picture or any such terminology. It is created by the predictions about the correlations in the visible outputs of certain conceivable experimental set-ups. ([4], p. 150)

The second part of the analysis, which unfolds the "difficulty ... created by the ... correlations," involves only very elementary mathematics. Clearly,

$$\text{Prob}\left(\{Z^1_{\mathbf{a}} = Z^1_{\mathbf{b}}\} \cup \{Z^1_{\mathbf{b}} = Z^1_{\mathbf{c}}\} \cup \{Z^1_{\mathbf{c}} = Z^1_{\mathbf{a}}\}\right) = 1\ ,$$

since at least two of the three (2-valued) variables $Z^1_{\boldsymbol{\alpha}}$ must have the same value. Hence, by elementary probability theory,

$$\text{Prob}\left(Z^1_{\mathbf{a}} = Z^1_{\mathbf{b}}\right) + \text{Prob}\left(Z^1_{\mathbf{b}} = Z^1_{\mathbf{c}}\right) + \text{Prob}\left(Z^1_{\mathbf{c}} = Z^1_{\mathbf{a}}\right) \geq 1,$$

and using the perfect anticorrelations (2) we have that

$$\text{Prob}\left(Z^1_{\mathbf{a}} = -Z^2_{\mathbf{b}}\right) + \text{Prob}\left(Z^1_{\mathbf{b}} = -Z^2_{\mathbf{c}}\right) + \text{Prob}\left(Z^1_{\mathbf{c}} = -Z^2_{\mathbf{a}}\right) \geq 1. \tag{4}$$

(4) is equivalent to the celebrated *Bell's inequality*. It is incompatible with (3). For example, when the angles between **a**, **b** and **c** are 120°, the 3 relevant quantum correlations $q_{\boldsymbol{\alpha\beta}}$ are all 1/4, implying a value of 3/4 for the left hand side of (4).

Let P be the hypothesis of the existence of noncontextual hidden variables for the EPRB-experiment, i.e., of preexisting values $Z^i_{\boldsymbol{\alpha}}$ for the spin components relevant to this experiment. Then Bell's nonlocality argument, just described, has the following structure:

$$\begin{array}{rccll} \text{Part 1:} & \text{quantum mechanics} + \text{locality} & \Rightarrow & \text{P} & (5)\\ \text{Part 2:} & \text{quantum mechanics} & \Rightarrow & \text{not P} & (6)\\ \text{Conclusion:} & \text{quantum mechanics} & \Rightarrow & \text{not locality} & (7) \end{array}$$

For this argument what is relevant about "quantum mechanics" is merely the predictions concerning experimental outcomes corresponding to (1–3) (with Part 1 using in fact only (2)). To fully grasp the argument it is important to appreciate that the content of P—what it actually expresses, namely the existence of the noncontextual hidden variables—is of little substantive importance for the argument. What is important is the fact that P is incompatible with the predictions of quantum theory.

The content of P is, however, of great historical significance: It is responsible for the misconception that Bell proved that (i) hidden variables are impossible, a belief until recently almost universally shared by physicists, and, more recently, for the view that Bell proved that (ii) hidden variables, while perhaps possible, must be nonlocal. Statement (i) is plainly wrong, since a hidden-variables theory exists and works, as mentioned earlier. Statement (ii) is correct, significant, but nonetheless rather misleading. It follows from (5) and (6) that *any* account of quantum phenomena must be nonlocal, not just any hidden-variables account. Bell's argument shows that nonlocality is implied by the predictions of standard quantum theory itself. Thus if nature is governed by these predictions, then *nature is nonlocal*. (That nature is so governed, even in the crucial EPRB-correlation experiments, has by now been established by a great many experiments, the most conclusive of which is perhaps that of Aspect [1].)

Concerning the wrongness of statement (i), some historical facts should be recalled. John von Neumann, one of the greatest mathematicians of the twentieth century, claimed to have mathematically proven that Einstein's dream, of a deterministic completion or reinterpretation of quantum theory (i.e., a hidden-variables theory), was mathematically impossible. Von Neumann's claim was almost universally accepted among physicists and philosophers of science. But Bohmian mechanics is a counterexample, so something had to be wrong with von Neumann's argument. Precisely what was wrong was elucidated by Bell in [3]. Nonetheless, many physicists continued to rely on von Neumann's proof and in recent years more commonly on Bell's inequality to support their rejection of the possibility of hidden variables.

Here's how Bell himself reacted upon learning of Bohmian mechanics ([4], p. 160):

> But in 1952 I saw the impossible done. It was in papers by David Bohm. Bohm showed explicitly how parameters could indeed be introduced, into nonrelativistic wave mechanics, with the help of which the indeterministic description could be transformed into a deterministic one. More importantly, in my opinion, the subjectivity of the orthodox version, the necessary reference to the 'observer,' could be eliminated. ...
>
> But why then had Born not told me of this "pilot wave"? If only to point out what was wrong with it? Why did von Neumann not consider it? More extraordinarily, why did people go on producing "impossibility" proofs, after 1952, and as recently as 1978? ... Why is the pilot wave picture ignored in text books? Should it not be taught, not as the only way, but as an antidote to the prevailing complacency? To show us that vagueness, subjectivity, and indeterminism, are not forced on us by experimental facts, but by deliberate theoretical choice?

In fact, Bell's examination of Bohmian mechanics led him to his nonlocality analysis. In the course of his investigation of Bohmian mechanics he observed that ([3], p. 452; [4], p. 11)

> in this theory an explicit causal mechanism exists whereby the disposition of one piece of apparatus affects the results obtained with a distant piece.
>
> Bohm of course was well aware of these features of his scheme, and has given them much attention. However, it must be stressed that, to the present writer's knowledge, there is no *proof* that *any* hidden variable account of quantum mechanics *must* have this extraordinary character. It would therefore be interesting, perhaps, to pursue some further "impossibility proofs," replacing the arbitrary axioms objected to above by some condition of locality, or of separability of distant systems.

In a footnote, Bell added that "Since the completion of this paper such a proof has been found." This proof was presented in his 1964 EPR-nonlocality paper [2] discussed here. (Because of publication delay the paper [3] was published after [2].)

Physicists' misconceptions notwithstanding, Bell did not establish the impossibility of a deterministic reformulation of quantum theory, nor did he ever claim to have done so. On the contrary, over the course of several decades, until his untimely death in 1990, Bell was the prime proponent, for a good part of this period almost the sole proponent, of the very theory, Bohmian mechanics, that he is supposed to have demolished.

References

[1] A. Aspect, J. Dalibard, and G. Roger. Experimental Test of Bell's Inequalities using Time-Varying Analyzers. *Phys. Rev. Lett.* **49**: 1804–1807, 1982.

[2] J. S. Bell. On the Einstein-Podolsky-Rosen Paradox. *Physics*, **1**: 195–200, 1964. Reprinted in [4].

[3] J. S. Bell. On the Problem of Hidden Variables in Quantum Mechanics. *Reviews of Modern Physics*, **38**: 447–452, 1966. Reprinted in [4].

[4] J. S. Bell. *Speakable and unspeakable in quantum mechanics.* Cambridge University Press, Cambridge, 1987.

[5] D. Bohm. *Quantum Theory.* Prentice-Hall, Englewood Cliffs, N.J., 1951.

[6] A. Einstein, B. Podolsky, and N. Rosen. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Phys. Rev.* **47**: 777-780, 1935.