Activitas Nervosa Superior (2019) 61:65–69
https://doi.org/10.1007/s41470-019-00046-z

ORIGINAL ARTICLE

Non-local Character of Quantum Mechanics: 20 Years Later

Tomasz Bigaj[1]

Received: 8 September 2018 / Accepted: 28 March 2019 / Published online: 16 April 2019
© The Author(s) 2019

Abstract
In this tribute to Henry P. Stapp, I outline the historical development of the problem of non-locality in quantum mechanics starting from the EPR argument through Bell's theorem up to Stapp's own contributions. My focus is on Stapp's 1997 proof of quantum non-locality that uses the logic of counterfactuals. On the basis of my own version of the proof presented in 2010, I give its informal analysis, emphasizing the role of Einstein's criterion of reality in the derivation. I conclude that the proof achieves its objective, and I briefly point at its significance for the philosophical mind-body problem.

Keywords Quantum mechanics · Non-locality · Bell's theorem · Counterfactuals · Criterion of reality · Mind-body problem

Introduction

The problem of non-locality in quantum mechanics came into view as a side effect of Albert Einstein's desperate attempts to reverse the direction which the new theory of subatomic phenomena had been taking. Einstein was particularly unhappy with the irreducibly probabilistic character of quantum mechanics and the inescapable limitations on our knowledge of quantum systems and their properties. In early discussions on the nature of quantum phenomena, the concept of unavoidable disturbances caused by the act of observation took center stage. Numerous thought experiments devised by Niels Bohr and Werner Heisenberg aimed at showing that each time we attempt to precisely measure the value of one parameter (e.g., position), the physical interaction with the quantum system which is a necessary part of the measurement process ineluctably disturbs the system causing it to lose the well-defined properties that it might have possessed beforehand (e.g., momentum). Thinking about these arguments, Einstein had a stroke of genius. What if it was possible to gain physical information about a system without actually physically interacting with it? But how can this be done? Here Einstein turned to the unique non-classical feature of the quantum-mechanical formalism which enables us to introduce joint states of many particles that display precise correlations between individual components—the so-called entangled states. In an entangled state, it is possible to infer the value of a given parameter for one particle by measuring this parameter on the other particle. Now, Einstein reasoned, if the particles were sufficiently far away from each other, there is no way a local measurement could disturb the state of the distant particle. This assumption gave rise to the famous EPR argument.

Einstein with Nathan Podolsky and Boris Rosen showed in their argument (Einstein et al. 1935) that if quantum mechanics is complete (i.e., there are no further hidden values beyond the standard probabilistic state), then there must occur a bizarre action at a distance violating the principles of relativity (no superluminal signaling). Deeming this conclusion absurd, they opted for the rejection of the assumption of completeness. In his response to the EPR argument, Bohr questioned the "mechanical" character of the disturbance of one system caused by a measurement of the other system, but his explanations were rather enigmatic (Bohr 1935). The whole matter lay relatively dormant until the 1960s, when a young physicist from CERN by the name of John Bell took the issue to a whole new level. Bell put together Einstein's thesis of incompleteness (in the form of a hidden-variable hypothesis) and a variant of the locality assumption which prescribes that a local measurement cannot change the value of a parameter characterizing a distant system, and he derived from this a certain inequality that can be subsequently showed to be violated by quantum mechanics (Bell 1964). Thus, the lesson from Bell's theorem appears to be that, contrary to what Einstein hoped

---

✉ Tomasz Bigaj
tfbigaj@gmail.com; t.f.bigaj@uw.edu.pl

[1] Institute of Philosophy, University of Warsaw, Warsaw, Poland

for, the hidden-variable hypothesis cannot be reconciled with the idea of locality.

However, the story does not end here. In a generalized variant of his theorem, Bell ventures to show that locality in one form or another should be rejected regardless of the issue of hidden variables (Bell 1975, 1981). As a matter of fact, his proof uses the notion of a "complete" physical state λ, which is customarily interpreted as containing some hidden values that go beyond the ordinary quantum-mechanical description. However, the proof goes through even if we assume that λ is just the standard quantum-mechanical state. The main assumption used in the derivation is the factorizability condition, which states that the joint probabilities of outcomes obtained on two particles in an entangled state can be rewritten as the product of individual probabilities, if we conditionalize all probabilities on the complete state λ. The factorizability is supposed to reflect the assumption of independence between both systems, and thus also the assumption of locality. From the condition of factorizability, Bell was able to derive the very same inequality that was known to contradict quantum-mechanical predictions.

The moral from the generalized Bell theorem regarding the issue of non-locality is somewhat tricky to draw. First off, it is a well-known fact that the condition of factorizability is violated in standard quantum mechanics (i.e., when we identify λ with the quantum-mechanical state of the particles). Thus, it may be surmised that Bell's theorem already presupposes that standard quantum mechanics is non-local and then proceeds to show that no addition of further hidden variables can turn it into a local theory. But this conclusion overlooks the fact that the exact meaning of the factorizability condition and its relation to the principle of no superluminal causation is far from obvious. Jon Jarrett in his 1984 paper (Jarrett 1984) showed that factorizability is equivalent to the conjunction of two further conditions which are commonly referred to as parameter independence and outcome independence (in Abner Shimony's terminology). Parameter independence expresses the statistical independence of the outcome obtained in one location from the selection of the observable to measure in the other location, while outcome independence secures the lack of statistical correlations between spatially separated outcomes. It can be easily showed that quantum-mechanical predictions regarding entangled states falsify outcome independence. However, this fact does not immediately prove that locality in the sense used in the EPR argument is violated. Einstein's understanding of the principle of locality was closer in spirit to parameter independence, as he assumed that *performing a measurement* (and not *selecting its outcome*) in one location should not have an instantaneous effect on the distant physical state. And Bell's theorem does not prove (at least not without further assumptions) that quantum-mechanical predictions necessarily invalidate *this* principle.

The Counterfactual Approach to Non-locality

The next important step in the development of the idea of non-locality in quantum mechanics has been taken by Henry P. Stapp. In an attempt to get rid of any extraneous assumptions in Bell-like arguments over and above the assumption of locality, Stapp turned to *counterfactuality*. The idea seemed straightforward enough: instead of talking about predetermined, objectively existing values of various observables, we can choose to talk hypothetically about outcomes that *would* be revealed if these observables were selected and measurements performed. It is possible that by using appropriate rules of counterfactual reasoning we could achieve the same conclusion as in the original Bell argument based on the hidden-variable hypothesis. Unfortunately, the logic of counterfactuals (a variant of modal logic) is notoriously tricky, and it is easy to fall prey to numerous fallacies when reasoning hypothetically about what would or could or might happen.

Since 1971, Stapp has produced a veritable cornucopia of counterfactual arguments—some rather informal, some highly formalized and technical—whose goal was to prove that the assumption of locality itself conflicts with the precepts of standard quantum mechanics.[1] And virtually each of his arguments has met with a barrage of criticism from physicists, philosophers, and logicians. The charges ranged from committing logical errors to smuggling the assumption of determinism (realism of possessed values). Unfazed by this hostile reception, Stapp continued to improve his counterfactual deductions, until in the 1997 paper "Non-local character of quantum mechanics" (Stapp 1997) he presented his ultimate word in the form of an argument based on the so-called Hardy case.[2]

As one of the critics of Stapp's arguments (coming relatively late to the field from the area of philosophy and logic with some background in physics), I was initially skeptical as to the success of his latest attempt to counterfactually prove quantum non-locality. Yet, the more I thought about Stapp's Hardy-based reasoning, the more I realized that it may actually be onto something. I decided to tinker a bit with the logical structure of the argument, replacing two independent locality assumptions used originally by Stapp with one all-encompassing condition, and straightening out some logical wrinkles, and I came up with a formal derivation in 16 rather tedious but overall straightforward steps (cf. Bigaj 2010, p. 64). The only glitch in the argument was that it actually required one premise over and above the principle of locality—

---

[1] (Stapp 1971) is the paper in which the idea of a counterfactual version of Bell's theorem occurred for the first time. I have analyzed extensively the 1971 argument as well as his later counterfactual derivations in (Bigaj 2006).

[2] (Stapp 2004) contains a condensed version of this argument with some minor logical errors eliminated. Stapp continued to produce various arguments in favor of the non-local character of quantum mechanics after that date (including an argument presented in his latest 2017 book, pp. 83–96), but to my knowledge they never made an extensive use of counterfactual reasoning.

namely the famous criterion of reality used originally by Einstein in his EPR argument. That sounded a bit disconcerting, as the possibility of rescuing locality by rejecting Einstein's criterion remains a logical option. Still, the criterion of reality is much weaker than the hidden-variable hypothesis, so progress has been made.[3] Below I present an informal rendering of the argument for the reader to appreciate the beauty and simplicity of Stapp's original idea.

The Hardy Case

The experimental setup in the Hardy case consists of a two-particle system whose components are spatially separated, so that no subluminal signal can travel from one particle to the other (for further details see Hardy 1992). Each particle can undergo measurements of two distinct observables ($L1$ or $L2$ for the left-hand side particle, and $R1$ or $R2$ for the right-hand side particle), and each measurement can yield one of two possible results (+ or –). The system has been prepared in a particular quantum entangled state which entails precise correlations between outcomes of spatially separated measurements. In particular, it is determined that if $L1$ and $R2$ are measured, and the first measurement yields "–", the measurement of $R2$ must show "+". The second prediction is that the outcome $R2$+ implies that the measurement of $L2$ must produce value "+". And the third relation in turn is that the outcome $L2$+ guarantees that observable $R1$, if selected for measurement, will admit value "–". Putting these three predictions together in the form of the chain of implications $L1- \Rightarrow R2+ \Rightarrow L2+ \Rightarrow R1-$, we may be tempted to infer, by the logical principle of transitivity, that the outcome $L1$– ensures that measuring $R1$ will yield outcome "–". Yet this is incorrect. The fourth quantum-mechanical prediction derived from the exact form of the Hardy state is that when $L1$ shows result "–", there is a *non-zero probability* that the measurement of $R1$ will reveal outcome "+".

The apparent failure of transitivity is explained away by noticing that it is impossible to simultaneously perform alternative measurements of incompatible observables $L1$ and $L2$ (or $R1$ and $R2$). However, we may attempt to reconstruct the transition from the occurrence of the outcome $L1$– to the necessary occurrence of the outcome $R1$–, using counterfactual reasoning. The main guiding principle in the derivation will be the condition of locality, understood as follows: making a free choice of an observable to measure at a given location X should not affect the physical situation in regions space-like separated from X. Or, to put it differently, if I had chosen observable A rather than B to be measured at X, all facts pertaining to locations space-like separated from X would have remained the same. Armed with this principle, we can make the first step of the argument by assuming that in reality observables $L1$ and $R1$ have been selected, with the measurement of the first observable revealing "–". We will attempt to prove that the principle of locality together with the three abovementioned conditionals imply that the outcome of the measurement of $R1$ must be "–" with probability one, thus violating the quantum-mechanical rules.

Given the actually obtained outcome $L1$–, we can make the following counterfactual prediction regarding the right-hand side particle: had we decided to measure $R2$ rather than $R1$, the outcome would have been guaranteed to be "+" (note that to get that we have to make the first use of the principle of locality which ensures that changing the measured observable from $R1$ to $R2$ does not alter the situation "$L1$–" in the left side of the system). And now we can appeal to Einstein's criterion of reality to argue that this prediction must be grounded in an objective property of the right-hand side system (let us symbolize this property as $\rho(R)$). The existence of this property is independent of whether we measure $R1$ or $R2$, as long as we can predict with certainty the outcome of $R2$ without disturbing the system. Moreover, due to the principle of locality, the existence of property $\rho(R)$ should also be independent of the choice of observable in the left-hand side system. Thus, we can now consider an alternative scenario (a "possible world") in which $L2$ has been selected instead of $L1$, and in this scenario, both $\rho(R)$ and $R1$ should still be present on the right-hand side.

Given that $\rho(R)$ describes what happens when $R2$ is measured, it is possible to appeal now to the second of the abovementioned conditionals and conclude that $L2$ (if selected) must produce outcome "+". Again, in order to prove that, the principle of locality turns out to be necessary. Suppose that $L2$ revealed value "–". Under the counterfactual assumption that $R2$ instead of $R1$ had been measured, the value $L2$– would have remained the same, but because the existence of $\rho(R)$ guarantees that $R2$ reveals "+", we have a violation of the second conditional $R2+ \Rightarrow L2+$. Thus the measurement of $L2$ must yield the result "+". Now the third conditional immediately secures that $R1$ produces the outcome "–", and this fact carries over to the original situation in which $L1$ is measured on the left-hand side (courtesy of the principle of locality, of course). Hence, we have arrived at the required conclusion $R1$–.

Conclusion

Let us take stock. We have been able to produce an argument that uses three quantum-mechanical predictions regarding the Hardy case, several instances of the counterfactual principle of locality, and one instance of Einstein's criterion of reality. This argument has led us to a consequence that violates the fourth quantum-mechanical prediction. This leaves us with the almost inevitable conclusion that the principle of locality must be false.

---

[3] Nowadays I believe that rejecting Einstein's criterion is a desperate move with nothing to gain, and therefore I concede that Stapp's argument has finally reached its objective. I expressed a different opinion, though, in (Bigaj 2006, p. 180). I admit that I have changed my mind since then.

The "almost" caveat refers to the slim possibility, mentioned earlier, that Einstein's criterion of reality may turn out to be the culprit. If we denied that the prediction regarding the outcome of the $R2$ measurement must be undergirded by an objective element of reality pertaining to the R system, we could insist that changing the L-observable from $L1$ to $L2$ may result in the invalidation of the above prediction without infringing the principle of locality. But how can we reasonably reject the existence of the property that makes the considered prediction true? Many influential philosophers of physics see Einstein's criterion as close to being analytically true and thus virtually unassailable.[4] Perhaps, instead of simply denying its existence, we could claim that the property characterizes not the right-hand side particle but the entire system, which would explain how it could be erased by the change in the left-hand side measurement. But this seems implausible. The property $\rho(R)$ informs us about the *local* outcome of a *local* measurement. If somehow this property manages to have its ontological "base" outside of this local system, this by itself should be seen as a breach of locality.

Which brings us to the real issue here, which is non-locality. The most straightforward way to avoid the contradiction is to admit, contrary to what we have assumed throughout the argument, that the experimenter's choice regarding the local observable *can* have an instantaneous impact on the physical situation in distant locations. That is, when we make an alternative choice of an experimental setup in one wing of the apparatus, the situation in the other wing cannot be assumed to remain intact. Based on the general argument given above, we cannot tell whether the non-local influence affects the outcomes revealed in the distant wing or perhaps the objective element of reality (the measurable property underlying the prediction regarding the $R2$ measurement) present there. But we can say for sure that what triggers this non-local influence is a free act of the observer, and not a particular response of the physical system under measurement.[5]

To make this point sharper, we can use John von Neumann's distinction between Process 1 and Process 2, also adopted by Stapp in his Realistically Interpreted Orthodox Quantum Mechanics (RIOQM). Process 2 involves interactions between physical systems only and is of no interest to us now. Process 1, on the other hand, introduces the notion of a conscious observer who interacts with a physical system. What is important is that Process 1 consists of two phases: a personal, subjective choice of the observer to probe a particular physical property, and a probabilistic elimination of all but one possible answers to the probe executed by nature (see Stapp 2017, p. 8). In our case, the first phase can be for instance the selection of observable $L1$ on the left-hand side of the system, while the second phase may result in the selection of value "–" and elimination of the alternative value "+". Now, the lesson from the Hardy case is that there exists non-local action at a distance that is initiated not by the second, but the first stage of Process 1. A conscious decision of the observer has physical consequences that reach beyond their immediate surroundings into faraway regions. Seen from the perspective of RIOQM, the non-local interactions revealed to exist in the quantum realm are not purely physical affairs but rather instances of the connections between the mental and the physical. This prompts interesting questions as to the exact nature of the phenomenon of non-locality discovered in quantum mechanics. In a sense, Einstein may be vindicated, since we have not proven the existence of a purely physical process that could propagate faster than light. Instead, it is the mental activity of a conscious observer that is responsible for physical changes in distant regions of spacetime. Whether the existence of such influences can teach us something about the nature of mind and consciousness (perhaps minds are not spatially limited to the physical bodies they occupy) remains to be seen. But one thing is for sure: there are still a lot of exciting questions to be asked at the border between quantum physics and philosophy.

Compliance with Ethical Standards

Conflict of Interest The author declares no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Bell, J. S. (1964), On the Einstein-Podolsky-Rosen paradox, *Physics* 1, 195–200. Reprinted in: J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, Cambridge: Cambridge University Press, 14–21.

Bell, J. S. (1975), The theory of local beables, TH-253-CERN. Reprinted in: J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, Cambridge: Cambridge University Press, 52–62.

Bell, J. S. (1981), Bertlmann's socks and the nature of reality. *Journal de Physique* 42, 241–261. Reprinted in: J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, Cambridge: Cambridge University Press, 139–157.

Bigaj, T. (2006). *Non-locality and possible words. A counterfactual perspective on quantum entanglement*. Frankfurt: Ontos Verlag.

---

[4] Tim Maudlin is one of those philosophers (private communication).

[5] One may complain that all this is old news, since the original EPR argument has already established that quantum mechanics without hidden variables must be non-local. To that it may be replied that the Hardy-case argument presented above is much more general. In fact, what we have demonstrated is that *every* theory that produces the correct experimental predictions in the Hardy case (which are borne out by experience) must be non-local. We do not have to assume any additional theses regarding measurements, collapses of the wave function and so on. In a sense the EPR argument indicates that standard quantum mechanics with collapse is non-local, while the current argument proves, beyond a reasonable doubt, that *reality* itself is non-local.

Bigaj, T. (2010). How to (properly) strengthen Bell's theorem using counterfactuals. *Studies in History and Philosophy of Modern Physics, 41*, 58–66.

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

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

Hardy, L. (1992). Quantum mechanics, local realistic theories and Lorentz-invariant realistic theories. *Physical Review Letters, 68*, 2981–2984.

Jarrett, J. (1984). On the physical significance of the locality condition in the Bell arguments. *Noûs, 18*, 569–589.

Stapp, H. P. (1971). S-matrix interpretation of quantum theory. *Physical Review D, 3*, 1303–1320.

Stapp, H. P. (1997). Non-local character of quantum theory. *American Journal of Physics, 65*, 300–304.

Stapp, H. P. (2004). A Bell-type theorem without hidden variables. *American Journal of Physics, 72*, 30–33.

Stapp, H. P. (2017). *Quantum theory and free will. How mental intentions translate into bodily actions*. Springer.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.