Quantum non-locality - It ain't necessarily so...

Żukowski and Brukner

Abstract

Bell's theorem is 50 years old. Still there is a controversy about its impli-
cations. Much of it has its roots in confusion regarding the premises from
which the theorem can be derived. Some claim that a derivation of Bell's
inequalities requires just locality assumption, and nothing more. Violations
of the inequalities are then interpreted as "nonlocality" or "quantum nonlo-
cality". We show that such claims are unfounded and that every derivation
of Bell's inequalities requires a premise — in addition to locality and freedom
of choice — which is either assumed tacitly, or unconsciously, or is embedded
in a single compound condition (like Bell's "local causality"). The premise
is equivalent to the assumption of existence of additional variables which do
not appear in the quantum formalism (in form of determinism, or joint prob-
ability for outcomes of all conceivable measurements, or "additional causes",
or "hidden variables", "complete description of the state" or counterfactual
definiteness, etc.). A certain irony is that perhaps the main message of vi-
olation of Bell's inequalities is that our notion of locality should be based
on an operationally well-defined no-signalling condition, rather than on local
causality.

The terms 'nonlocality' or 'quantum non-locality' are buzzwords in founda-
tions of quantum mechanics and quantum information. Most of scientists
treat these terms as a more handy expression *equivalent* to the clumsy "vio-
lation of Bell?s inequalities". Unfortunately, some treat them seriously. Even
more unfortunately Bell himself used such terms in later works [1, 2] [26].

There are two approaches to Bell?s theorem which we would like to critically
analyze:

- (A) Derivation of a Bell inequality relies only on the premise of locality,
  and nothing more.

- (B) Derivation of Bell inequality does rely on the additional premise
  of determinism, or any of the notions listed in the abstract. However,
  these notions can be derived from the premises of locality, the freedom
  of an experimenter to choose the setting of his/her local apparatus and
  quantum predictions, and nothing more.

Consequently, violations of Bell's inequalities in either approach is inter-
preted as a demonstration of "non-locality". (As it is widely accepted that it

is difficult to dismiss the 'freedom of choice' assumption this aspect of Bell's theorem will be not discussed.) Supporters of (A), state that it is not clear what realism, in the context of Bell's theorem, should mean, and that those who use this notion most probably confuse it with determinism — the notion that outcomes for all possible measurements are pre-determined. Additionally, they claim that determinism, or the other versions of the assumption listed in the abstract, needs not to be assumed neither in stating nor in proving Bell's inequalities, and thus cannot be the issue. Followers of (B) see this differently. They use pre-determined values in derivation of Bell's inequalities but claim that it is an implication of locality, freedom of choice, and experimentally observed perfect correlations. Hence, they basically follow the Einstein-Podolsky-Rosen (EPR) argument [3].

We show that Bell's condition of "local causality" [4] which proponents of view (A) often adopt is equivalent to the assumption of existence of a (positive and normalized) joint probability distribution for the values of all possible measurements that could be performed on an individual system, no matter whether any measurement — and which measurement — is actually performed. This shows that local causality is a *compound* condition, which assumes the existence of causes which are not present in quantum mechanics, and that the causes act locally (i.e., their influence is bound to the future light cone). Hence, *locality is not the sole assumption* of view (A) based on local causality. This refutes the view (A).

We refute the view (B): any attempts to derive determinism, or hidden-variables, via an EPR argument [3] are futile. The argument of EPR was based on an assumption, which is often not noticed or forgotten: *counterfactual definiteness* [5, 6]. This assumption allows one to assume the definiteness of the results of measurements, which were actually *not* performed on a given individual system. They are treated as unknown, but in principle *defined* values. This is in a striking disagreement with quantum mechanics, and complementarity principle. Therefore, it cannot be glossed over as something not worth mentioning as an additional assumption.

Our conclusion is that, despite suggestions to the contrary, one always assumes an additional premise, except locality and freedom, to derive any Bell inequality. The additional premise is either assumed tacitly, or unconsciously[27], or is embedded in a single compound condition (like Bell's

'local causality'), from which the inequalities are derived.

1. Locality and Nothing More?

Bell's discussion of the premises behind his theorem, which differs substantially with respect to his original approach is given most extensively in his final published paper 'La nouvelle cuisine' [4]. We will base our discussion on it, as the work is often referred to by advocates of (A). In this work Bell fully defines his new notion of local causality.

Consider two space-like separated parties, Alice and Bob. Either of them has a choice between a number of measurement settings. Denote, respectively, by x and y Alice's and Bob's choice of measurement settings and by A and B their outcomes. Correlations of the outcomes are described by a conditional probability distribution $p(A, B|x, y)$.

It is argued, that ??the situation described by $p(A, B|x, y)$ may arise out of a statistical mixture of different situations traditionally labeled by λ and sometimes called 'causes'. The probabilities therefore acquire the following form $p(A, B|x, y) = \int d\lambda \, \rho(\lambda) p(A, B|x, y, \lambda)$ where $\rho(\lambda)$ is a probability distribution. The properties of conditional probabilities allow one to put $p(A, B|x, y, \lambda) = p(A|B, x, y, \lambda) p(B|x, y, \lambda)$. Next, local causality stated by Bell as 'The direct causes (and effects) of [the] event are near by, and even the indirect causes (and effects) are no further away than permitted by the velocity of light', allows one to state that 'what happens on Alice's side does not depend on what happens on Bob's side' and vice versa [7]. This results in $p(A|B, x, y, \lambda) = p(A|x, \lambda)$ and $p(B|x, y, \lambda) = p(B|y, \lambda)$. Finally, we obtain the general mathematical structure underlying derivations of all Bell's inequalities:

$$p(A, B|x, y) = \int d\lambda \, \rho(\lambda) p(S|x, \lambda) p(B|y, \lambda) \tag{1}$$

Proponents of (A) then argue that besides 'locality' nothing more has been assumed in the derivation of equation (1). Obviously determinism is not required. The condition (1) holds also for any local stochastic (indeterministic, hidden variable) theory. While a deterministic theory specifies which outcome, exactly, will happen under a given λ, a stochastic theory specifies only the probabilities for various outcomes that might be realized. (Note

that the former is a special case of the latter, with the probabilities always being exclusively 1 or 0.). So, where, if at all, "realism" of any sort could be hidden in condition (1)?

Note first, that in the case of the mixed separable states, one may think that λ specifies the "actual" quantum state in the probabilistic mixture. All this would agree with the formula (1). However when the joint quantum state is an entangled one, and especially if it is additionally pure, and no additional λ's are introduced, one cannot have a factorization like (1), see e.g., Ref. [7], or appendix B. Thus the formula (1) is equivalent to the introduction of additional hidden parameters, λ's, which are *not* present in quantum theory. The λ, which enter Eq. (1), can pop up under many guises such as, e.g., 'the physical state of the systems as described by any possible future theory' [7], 'local beables', 'the real state of affairs', 'complete description of the state', etc. Since λ do not appear in quantum mechanics, thus they are (good old) *hidden variables.* Anything on which one conditions probabilities, which gives *different structure* to formulas for probabilities than quantum mechanical formalism[28], is a hidden variable *per se.* Bell himself writes 'λ denote any number of hypothetical additional variables needed to complete quantum mechanics in the way envisaged by EPR' ([4], page 242). This sentence of Bell's is often forgotten by supporters of (A). As we will see next, we can try to introduce λ's not existent in quantum theory, but this will result in a condition which is stronger than locality alone.

The condition (1) allows assignment of (positive and properly normalized) *joint* probabilities for the (local) values of the entire set of pairs of (local) measurements. Denote as A_x the value of the outcome A pertaining to the situation in which Alice chooses setting x, which, for example, in the Clauser-Horne-Shimony-Holt (CHSH) [8] scenario has two possible values, denoted here as 1, 2, and similarly define B_y ($y = 1, 2$). We can always introduce a joint probability

$$p(A_1, A_2, B_1, B_2 | \lambda) = p(A | x = 1, \lambda) p(A | x = 2, \lambda) p(B | y = 1, \lambda) p(B | y = 2, \lambda)$$
$$(2)$$

When the probabilities are 0 or 1 the model is deterministic, otherwise it is stochastic. Hence, we have defined a joint probability for all possible outcomes under all possible pairs of settings in the CHSH scenario. These outcomes include, for a single run of a Bell experiment, the actually measured

ones and on an equal footing the ones which could have been potentially measured. Note that λ is no m??ore necessary here, as to get predictions for such a model it is enough to know $p(A_1, A_2, B_1, B_2)$, which is given by $\int d\lambda\, p(A_1, A_2, B_1, B_2|\lambda)\rho(\lambda)$. Starting from equation (1) one can introduce the joint probability $p(A_1, A_2, ..., B_1, B_2, ...)$ for an arbitrary number of settings. Note that the existence of such joint probability implies the existence of $p(A_1, A_2, ...) \geq 0$, which has nothing to do with locality and is already in conflict with the Kochen-Specker theorem [9].

Conversely, starting from the existence of the joint probability one can derive (1) for any pair of local settings. This is because in Kolmogorovian probability theory, which is the standard axiomatization of classical probability, if Ω is a probability space with probability measure $rho(\lambda)$ and A is a measurable set contained in Ω, then indicator of A, that is $\Omega\ A$, gives by the formula $P(A) = \int_\Omega \chi_A \rho(\lambda) d\lambda$ the probability of the event A. The probability $p(A_1, A_2, ..., B_1, B_2)$ can be modelled by $p(A_1, A_2, ..., B_1, B_2|\lambda) = \chi_{A_1}(\lambda)\chi_{A_2}(\lambda)\chi_{B_1}(\lambda)\chi_{B_2}(\lambda)$, that is, one can put $p(A_1, A_2, ..., B_1, B_2) = \int_\Omega \chi_{A_1}(\lambda)\chi_{A_2}(\lambda)\chi_{B_1}(\lambda)\chi_{B_2}(\lambda)\rho(\lambda)d\lambda$. Then by calculating the marginals, i.e., summing up the probabilities over the outcomes of the observables which are *not* measured in an actual experiment, one obtains: $P(A, B|x, y) = \int_\Omega \chi_{A_x}(\lambda)\chi_{B_y}(\lambda)\rho(\lambda)d\lambda$, where $\chi_{A_x}(\lambda)$ and $\chi_{B_y}(\lambda)$ are the indicator functions for the events A_x and B_y respectively. This is of course, equivalent to equation (1).

Thus, we see that local causality condition (1) is mathematically equivalent to the assumption of joint probabilities, $p(A_1, A_2, ..., B_1, B_2)$. The latter *is* a form of realism: complementary observables are treated as mere numbers ('c-numbers' in Dirac's terminology [10]). One can speak about the joint probability of A_1, A_2 and B_1, B_2 only if in a theoretical construction these all values coexist together independently of which experiment is actually performed on either side, and in this sense are "real". Expressing it differently, the existence of a proper distribution $p(A_1, A_2, ..., B_1, B_2)$ means that we model everything with Kolmogorovian probabilities, which always have a lack of knowledge interpretation, with the underlying Ω is treated as a sample space. Indeed, according to the previous results due to Fine [11], Hall [12], Gill et al. [13] and others, the probabilities of a local stochastic model *can* always be understood as stemming from our ignorance about supposed local deterministic values. However, we do not need to insist on the existence of deterministic values of

A_1, A_2 and B_1, B_2. The supposed existence of their joint probabilities is sufficient for our argument. In fact, the entire discussion on local deterministic vs. stochastic models — on which the proponents of view (A) build up their argument — is irrelevant for the current discussion. This is most evident if one takes into account the Greenberger-Horne-Zeilinger (GHZ) [14] type of argument, which involves only perfect correlations between *three* (or more) systems and thus no stochastic model can ever recover them. (Note, that it is symptomatic, that followers of view (A) always base their argument on correlations between two systems.) Furthermore, assuming a fundamental local indeterminism (that is some fundamental finite irreducible stochasticity, which does not allow one to use deterministic models ? even as only a mathematical tool) would lead to different bounds of the Bell inequalities. We show this in Appendix A. Somewhat ironically, not any assumption of realism but its most stringent version, namely determinism itself, must be invoked to obtain the proper bounds of Bell's inequalities.

The following observation can be made. Some researchers are willing to accept that an outcome measured on a single (local) system may be intrinsically probabilistic. However, they do not accept the same for correlations between outcomes measured on several such systems; the probabilities for correlations are always taken to be reducible to probabilities for local outcomes in the form like (1). Confronted with the experimental violation of Bell's inequalities, they accept an inherently probabilistic explanation for an individual quantum system but hold fast to a pseudo-causal [29] nonlocal one for correlations (e.g., by introducing faster-then-light influences between distant quantum systems).

2. Derived Realism?

The school of thought (B) claims that one does need any version of the assumption of realism, or hidden variables to derive Bell's inequalities. However, it is purported, that everything is deductible via EPR type reasoning. The current classic expositions of such line of thought can be found in Refs. [15-17]. For example, the argumentation in Ref. [17] tries to oppose the results of the "Free Will Theorem" of Ref. [18]. To this end, the following two logical statements are both claimed to be true. The first one is, that the logic of Bell's theorem is in the validity of the following implication (we use the

following terminology for the structure of a logical implication: *assumption* ⇒ *thesis*):

$$freedom \text{ and } QF \text{ and } locality \Rightarrow contradiction \qquad (3)$$

where *freedom* stands for the freedom of choice of local measurement settings assumption, *locality* for locality assumption and *QF* for quantum formalism. On the other hand, the logic of the EPR paper supposedly allows to establish a yet another implication as valid:

$$freedom \text{ and } QF \text{ and } locality \Rightarrow determinism \qquad (4)$$

The validity of these statements is used to challenge the implication of Ref. [18], which as put in Ref. [17], is

$$freedom \text{ and } QF \text{ and } locality \text{ and } determinism \Rightarrow contradiction \qquad (5)$$

Obviously, if the implication (3) is valid, the assumption which leads to it cannot hold. This is because *contradiction* means a false statement (like $2 \geq 2\sqrt{2}$). The rules of Aristotelian two-valued logic say that if the thesis of a *valid* (that is, true) implication is false, then its assumption must be false too (*modus tollens*, see the truth table of logical implication). Thus the assumption of (4), as it is the same as of (3), must be false and one cannot 'determine' whether determinism is true, from the validity of the implication (4), as from false statements one can derive via a valid implication both true and false statements (see again the truth table).

Of course the first implication (3) is highly appealing to proponents of (A), while the second one (4) for supporters of (B). Alone the first implication would mean *non-locality* of quantum mechanics (provided freedom holds). But, "it ain't necessarily so" [19]. It is the implication (5), which correctly describes the situation, see below.

A. EPR Reasoning

The works [15] and [16] are based on an assumption of the validity of the EPR reasoning. For example in Ref. [16] one can find that 'the existence of these local, deterministic, non-contextual hidden variables [...] is not simply assumed, but is inferred from Locality plus a certain subset of the quantum mechanical predictions, using (in essence) the EPR argument.' The work [15] is aimed at showing that: 'The hidden variables posited by Bell are not

7

an 'ad hoc assumption' but, rather, a logical implication of locality.'

The reasoning of EPR is often presented in the form of (4), but this is wrong. The logical structure of the EPR argumentation about elements of reality is

$$\text{freedom and QF and locality and counterfactual definitness}$$
$$\Rightarrow \text{for specific observables elements of reality exist for the EPR state.} \quad (6)$$

The thesis of the implication is true, provided one does not make an unfounded generalization of it to *arbitrary* observables and *arbitrary* states. Indeed, for the EPR state and momentum and position observables (P and Q), elements of reality seem to be a consistent notion. However this is arrived at by considering two situations of which only one can be the case in the given run of the experiment (measuring either P and Q, page 780 of [3]). This is counterfactual definitness at work.

As a matter of fact such a generalization mentioned above is the effective claim of EPR, as they aimed to prove incompleteness of the entire theory of quantum mechanics. A missing (hidden) part of the theory would be according to them *inherently deterministic "elements of reality"*. By EPR, since they correspond to values of observables which can be predicted with certainty (i.e., with probability equal to unity), an act of measurement just displays the previously hidden values. However, there is a logical flaw in the paper of EPR. General existence of elements of reality is just their conjecture, based on just *one* example, while they claim that this is their thesis, holding always. Why conjecture? EPR introduced elements of reality just for the 'original EPR' state. And this was successful, that is it did not lead to a contradiction, only because they limited themselves to specific observables, which do not exhaust all possible ones. Note here, that for different observables and the original EPR state, in Ref. [21], one can find a proof of internal inconsistency of the EPR concepts.

In their conclusions EPR tacitly assumed that one can establish local elements of reality for all states with perfect correlations, and for all observables. This is wrong, as it can be directly shown in the case of the GHZ states. The GHZ reasoning shows that the very definition of elements of reality must be wrong, as it implies contradictory values (a $1 = -1$ contradiction in the case of GHZ correlations, or $2\sqrt{2} \leq 2$ in the case of the CHSH version of the Bell theorem). This means that the thought-provoking paper of EPR does not

contain a valid general statement on quantum mechanics! One could only *counterfactually* wonder what would have been the views of EPR, had the GHZ paper appeared before 1935.

Note further that the *thesis* of Bohm's version of the EPR implication [20]

$$\textit{freedom and QF and locality and counterfactual definitness}$$
$$\Rightarrow \textit{elements of reality exist for the two } - \textit{ spins } - \textit{ 1/2 } - \textit{ singlet state,}$$
(7)

is plainly *not* true. The thesis of this implication leads to a $2 \geq 2\sqrt{2}$ contradiction, if one forms the CHSH inequality for the elements of reality, and uses again quantum predictions. Of course, counterfactual definitness was a tacit assumption in Bohm's reasoning.

Comparing the relation of (6) with the views of proponents of (B), one sees the following differences with the simplified description of the EPR result given by (4). The assumption of *freedom* is sometimes thought to encompass also *counterfactual definitness*, and pre-determined values are thought to be derivable for all states with perfect correlations. The second is invalidated by the Bell theorem and even more strikingly by the GHZ reasoning. However, the basic problem here is that freedom does not encompass counterfactual definitness. This is evident, for example, in the case of the CHSH inequality where we discuss four situations (four possible pairs of local settings), only *one* of which can actually occur for the given pair or particles (run of the experiment). However, the proper bounds of inequalities (see appendix A) are derived by an algebraic manipulation of values for a *single pair*, for the actual situation and three counterfactual ones ("had one or both the observers chosen different settings")[30]. The results that would have been obtained in such cases are then treated as unknown, but nevertheless definite real numbers (± 1 in the case of the CHSH scenario). This is counterfactual definitness *per se*. In this way an effective determinism enters.

Counterfactual definiteness is directly irreconcilable with Bohr's complementarity principle. The principle says that, once a value of an observable is measured, one is not allowed to even speak about values for complementary observables. This is reflected by quantum formalism, according to which probability distributions for complementary (non-commuting) observables are in general not defined, or rather they a *undefinable* within the theory.

9

Thus, introduction of hidden variables, determinism, counterfactual definiteness etc., is not a minor point, a soft option, or something so obvious that may be treated as a tacit, indisputable assumption. It goes directly against the very essence of quantum mechanics.

3. Conclusions

The terms 'nonlocality' or 'quantum nonlocality' suggest that there is some 'spooky' faster-than-light influence between distant quantum systems (for a critical analysis see Ref. [12]). While the possibility of such an influence cannot be excluded, it is just a one-sided view. A failure of realism in all of its forms, for example, futility of considering joint probabilities for outcomes for the entire set of conceivable measurements or rejection of counterfactual reasoning, is an equally valid option, just as a failure of both realism *and* locality.

In "La Nouvelle Cuisine" Bell describes failure of *local causality* in the quantum world. This failure does not mean that we have to accept non-local causality. Individual events may have *spontaneous, acausal* nature. There seem to be no need to go beyond quantum mechanics within this aspect [31]. Paradoxically, it was Einstein who reluctantly introduced the notion of spontaneous events, which might be after all the root of Bell's theorem. The lesson for future could however be that we should build the notion of locality on the operationally clear "no-signalling" condition — the impossibility to transfer information faster-than-light. After all this is all what theory of relativity requires.

The moral of the story is that Bell's theorem, in all its forms, tells us *not* what quantum mechanics is, but what quantum mechanics *is not*.

4. Appendix A

We show that inherently stochastic local hidden variable theories (that is, the theories with some fundamental finite irreducible stochasticity) lead to lower bounds than the standard ones of Bell inequalities. To this end we derive a Bell inequality for a local *strictly* stochastic theory. The latter is equivalent to the requirement that all $p(A|x, \lambda)$ and $p(B|y, \lambda)$ are strictly different from

0 or 1. Therefore, suppose that there is inherent stochasticity parameter, $s > 0$, which gives the following upper bound: for any λ

$$s < p_s(X|z, \lambda) < 1 - s \qquad (8)$$

where $X = A, B$, and $z = x, y$. As Bell's inequalities are in form of multi-linear combinations (functions) of the underlying probabilities, their bounds (maxima and minima) are at the border of the region of validity of the probabilities. In the standard case the bo??rders are $0 \leq p(X|z, \lambda) \leq 1$. Thus, the standard bounds cannot be reached in the case of probabilities satisfying (8). For example, take the CHSH inequality [8]. With (8), for every λ one can introduce the expectation value $I(z, \lambda) = \sum_{x=\pm 1} X p(X|z, \lambda)$ (we assume here the usual spectrum of the Bell observables: $X = \pm 1$). Obviously $-1 + 2s \leq I(z, \lambda) \leq 1 - 2s$. Thus, 'renormalized' variables $I' = \frac{1}{1-2s} I$ give the usual of the CHSH inequality, ?that is 2. However, we have

$$\langle I(x, \lambda)I(y, \lambda) + I(x, \lambda)I(y', \lambda) + I(x', \lambda)I(y, \lambda) - I(x', \lambda)I(y', \lambda) \rangle \leq 2(1-2s)^2 < 2$$

The bound[32] of the CHSH inequality in a local (intrinsically) stochastic theory is strictly smaller than 2. One must assume *both* locality and determinism in order to obtain the proper bounds for Bell's inequalities[33].

5. Appendix B: Alternative Analysis

One can formulate the analysis of the assumption of local causality also in another alternative way. The principle of local causality, $p(A|B, x, y, \lambda) = p(A|x, \lambda)$ and $P(B|A, x, y, \lambda) = p(B|x, \lambda)$, implies in *the case of correlated systems*:

- An easily derivable relation

$$p(A|x, y, \lambda) = p(A|x, \lambda) \qquad (9)$$

 which is a form of the *no-signaling condition*. This is because under lo-cal causality $p(A|x, y, \lambda) = \sum_B P(A, B|x, y, \lambda) = p(A|x, \lambda) \sum_B p(B|y, \lambda)$.

- Existence of *at least two different* values of 'causes' λ, which makes λ a non-trivial parameter (variable) outside quantum formalism (i.e., a non-trivial hidden variable).

11

Why the second implication? If one Assumes just *one* common cause and local causality, this immediately leads to *no* correlations. To see this consider the following relations:

$$P(A, B|x, y, \psi) = P(A|B, x, y, \psi)P(B|y, \psi) = P(A|x, \psi)p(B|y, \psi) \qquad (10)$$

where ψ is the sole cause, which can be thought of as a pure (entangled) quantum state describing the preparation. Here we used first probability rules and later local causality. The factorization means: *no* correlations. In order to describe correlations one *must* have at least two values for λ and this implies that λ is outside of the quantum formalism, as the only common cause, in the case of pure quantum states, allowed by quantum mechanics ψ is the quantum state. Just one $\lambda = \psi$ cannot give us any correlations whatsoever.

Thus local causal theories (of correlated systems) are a subset of non-trivial hidden variable theories (ones allowing for λ, outside of the quantum formalism), and local theories (which additionally impose $p(A|x, y, \lambda) = p(A|x, \lambda)$. That is, a local causal theory is a specific example of a local hidden variable theory.

Quantum mechanics (QM) and quantum field theory (QFT) are no-signalling. In QFT it is assumed the for two gauge invariant observables, O_1, O_2 at two spatially separated space time points ξ_1, ξ_2 one has $[O_1(\xi_1), O_2(\xi_2)] = 0$, and in QM one has for the results of a local observable x defined on one of the local subsystems: $P(A|x, y, \lambda) = P(A|x, \lambda)$ (the meaning of the symbols is as above). Note that both QFT and QM are *not* local causal, but this is precisely the point of Bell's theorem. Briefly, if one wants to work outside local hidden variable theories, one cannot even formulate the local causality principle. Local causal hidden variable theories assume first of all non-trivial hidden variables λ, and next that a form of locality principle is satisfied, in the form of $P(A|B, x, y, \lambda) = P(A|x, \lambda)$, etc.

6. Acknowlefgements

CoQuS, SFB FoQuS, and the Individual Project 2462.

References

[1] J. S. Bell, Atomic cascade photons and quantum-mechanical nonlocality, Comments on Atomic and Molecular Physics 9, 121 (1980).

[2] J. S. Bell, The theory of beables, TH-2053-CERN (1975).

[3] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. 47, 777 (1935).

[4] J. S. Bell, La nouvelle cuisine, in Speakable and Unspeakable in Quantum Mechanics, 2nd ed., Cambridge University Press, 2004.

[5] W. M. de Muynck, W. De Baere, Found. Phys. Lett., 3 325 (1990).

[6] W. M. de Muynck, W. De Baere, H. Martens, Found. Phys. 24, 1589-1664 (1994).

[7] N. Gisin, Found. Phys., 42, 80 (2012)

[8] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. 23, 880 (1969).

[9] S. Kochen, and E. Specker, J. Math. Mech. 17, 59 (1967).

[10] P.A.M. Dirac, The Principles of Quantum Mechanics, The Clarendon Press, Oxford (1930).

[11] A. Fine, Phys. Rev. Lett. 48 (1982) 291. A. Fine, J. Math. Phys. 23 (1982) 1306.

[12] M. J. W. Hall, arXiv:0909.0015v3; M. J. W. Hall, Int. J. Theoret. Phys. 27 (1988).

[13] R. D. Gill, G. Weihs, A. Zeilinger, M. Zukowski, Europhys. Lett. (2003) 61, 282-283.

[14] D. M. Greenberger, M. A. Horne, and A. Zeilinger, in Bell?s Theorem, Quantum Theory, and Conceptions of the Universe, edited by M. Kafatos (Kluwer Academic, Dordrecht, 1989); D. M. Greenberger, M. A. Horne, A. Shimony, and A. Zeilinger, 1990, Am. J. Phys. 58, 1131 (1990).

[15] T. Norsen, Found. Phys. Lett. 19, 633 (2006).

[16] T. Norsen, Against Realism, Found. Phys. 37(3), 311-340 (2007).

[17] R. Tumulka, Found. Phys. 37, 186 (2007).

[18] J. Conway and S. Kochen, Found. Phys. 36, 1441 (2006). quant-ph/0604079.

[19] George Gershwin and Ira Gershwin 1935 Porgy and Bess (the opera).

[20] D. Bohm, Quantum Theory (Prentice-Hall, Englewood Cliffs, NJ, 1951).

[21] K. Banaszek, K. Wodkiewicz, Phys. Rev. A 58, 4345 (1998).

[22] A. Stairs, unpublished, http://www.terpconnect.umd.edu/?stairs/papers/EPR Illusion.pdf.

[23] A. Peres Quantum Theory: Concepts and Methods (Kluwer Academic Publishers, New York, 2002).

[24] C. A. Fuchs and A Peres, Physics Today 53 (2000).

[25] A. Peres, Am. J. Phys. 46, 747 (1978).

[26] In Ref. [2] a chapter is called 'Locality inequality?' and it refers to the Bell-CHSH inequality. Consistently with this, in Ref. [1, 2] Bell writes that the famous inequalities imply non-locality. However, such statements are not in all works of Bell. Even in Ref. [1], in the last paragraph, he writes the following, as one of the logically possible implications of violations of the inequalities: 'The lesson of quantum mechanics is not to look behind the predictions of the formalism', and further '... it may be that Bohr's intuition was right — in that there is no reality below some 'classical' 'macroscopic' level'.

[27] 'Par ma foi! il y a plus de quarante ans que je dis de la prose sans que j'en susse rien ...' (Moliere, Le Bourgeois Gentilhomme, 1670)

[28] The basic quantum structure is $\text{Tr}\rho\Pi$, where ρ is the state and Π is a projector.

[29] That is, not obeying relativistic causality.

[30] For a more extended analysis see, e.g., the recent Ref. [22].

[31] In his masterpiece book [23], Peres, when discussing Bell's theorem, uses a phrase 'non-locality is inescapable'. However, a careful reader would notice that he firmly assumes that counterfactual reasoning is acceptable (statements like had we measured a different observable than the actual one which was measured, we would have obtained a value, say x). Only using a counterfactual reasoning measured and unmeasured, but potentially measurable, variables can be put on an equal footing. If one does not consider this assumption as disputable, non-locality is an inescapable consequence of Bell's theorem. In his later work with Fuchs [24], he rejects non-locality. Also his well known slogan 'unperformed experiments have no results' [25] is a clear rejection of counterfactual definitness.

[32] One can choose infinitesimally small but nonzero s such that the new bound is arbitrarily close to 2 and still we have, in the strict mathematical sense, a non-deterministic theory. However, the different viewpoints (here determinism vs. nondeterminism with an infinitesimal s) would not be operationally distinguishable.

[33] Note that in order to obtain the usual bound it is enough to assume that e.g., results of Bob are *deterministic*, and one of the results on Alice's side is deterministic too. Of course such an asymmetry of indeterminism is ridiculous. This is why we assumed an irreducible stochasticity, s, at both sides of the experiment, for all settings.