The Empty Scope of Bell’s Theorem
Horace Yuen

To cite this version:
Horace Yuen. The Empty Scope of Bell’s Theorem. 2019. hal-02161557

HAL Id: hal-02161557
https://hal.archives-ouvertes.fr/hal-02161557
Submitted on 20 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
The Empty Scope of Bell’s Theorem

Horace P. Yuen

Department of Electrical Engineering and Computer Science
Department of Physics and Astronomy
Northwestern University
Evanston, IL 60201
hyuen081@gmail.com

Abstract

Bell inequality is violated in quantum mechanics and in experiments. Bell’s theorem, on the other hand, concludes that quantum physics or nature is hence nonlocal. Bell’s derivation of such a theorem comes in two versions, BT1 in which he started with determinism assumption in his 1964 original paper, and BT2 in his 1976 paper where he started with a "local causality" condition (LC). In this paper we show that (LC) does not follow from the physical concept of local causality except in the determinism limit. A local hidden variable model for the singlet pair is given that yields the correct quantum predictions for pair correlations. We also show determinism flagrantly violates quantum predictions. Thus Bell’s theorem has no scope of applicability on possible hidden variable theory for quantum mechanics.
Bell inequality (BI) is a set of mathematical relations on the binary correlations [1] among some random variables. To avoid confusion, we distinguish for a given index set \( i \) the underlying physical variables \( \mathcal{A}_i \), the corresponding Hermitian operator observables \( \hat{A}_i \) in quantum mechanics, the random variables \( A_i \) when \( \hat{A}_i \) is measured in a quantum system, and the possible values \( a_i \) that may result which would be an eigenvalue of \( \hat{A}_i \). In a sequence of experiments with the same quantum state \( |\psi\rangle \), the relative frequency of each \( a_i \) is given by the probability \( |\langle a_i |\psi\rangle|^2 \) from the Born rule for nondegenerate \( \hat{A}_i \) spectrum. For the spin 1/2 singlet state \( |\psi\rangle = \frac{1}{2}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \) on two sites A and B, BI focuses on three or four pairs of correlations on four spin variables of the two sites. The original three-correlation BI was generalized to four correlations in [2], the CHSH inequality which we subsume under the term BI, for which experimental violation in accordance with quantum predictions has been repeatedly confirmed [3].

What are the assumptions used in the derivation of BI, one of which must not hold from the experimental violation of BI. In [1] the assumptions are taken to be "determinism" and "locality"; see [4] for discussions on historical terminology. Determinism means the hidden variables determine the values of all \( \mathcal{A}_i \) with probability 1 as in classical mechanics. "Locality" is a term used by Bell then and is misleading, and his other term for the same mathematical condition, "separability," is preferred [4]. Bell in [1] thought the violation of BI shows "locality" cannot be maintained in quantum physics, which is then nonlocal in the sense of "spooky action at a distance" [5, p. 143], a nonlocal influence of one spin measurement result on the other spin measurement result, in the absence of which BI would be satisfied instead. We will call such influence "Bell nonlocality" which, in this singlet correlation case, does not lead to superluminal signaling and hence may not violate special relativity badly. Bell nonlocality can be contrasted with "Einstein locality," which is taken to say one system does not affect another it does not interact with according to the known laws of physics. This conclusion of nonlocality from BI violation is called Bell's theorem, which is more than BI, and we would call this "BT1," the first version of Bell's theorem in the good distinction introduced in [4] to differentiate it with the second version "BT2" developed in [6] and Bell's subsequent papers by using the derivation of BI in [7].

Bell emphasized on occasion that determinism is not an assumption, but a consequence of "local realism" embodied in the EPR argument [8]. See, e.g., [5]. Thus the logic of Bell's
theorem is given by the following argument which modifies the one given in [9, p. 234]:

\[
\begin{align*}
\text{EPR} & \quad \text{locality} \quad \rightarrow X \\
\text{Bell} & \quad X \quad \rightarrow \text{BI} \\
\text{QM or Expt} & \quad Y \quad \rightarrow \text{not BI} \\
\text{Conclude} & \quad Y \quad \text{is nonlocal}
\end{align*}
\]

Here Y is quantum mechanics (QM, as a theory) or nature (from experiments), either of which shows BI is violated. Hence starting with Einstein locality one concludes nature and quantum mechanics is nonlocal. In BT1, X is determinism. In BT2, X is the mathematical consequence (LC) of the physical concept of "local causality" [10], [4]. This (LC) condition is equivalent to the factorization condition first introduced in [7] for deriving BI.

Bell himself favored BT2 and he always discussed it in lieu of BT1 after 1976. As thought by many others and discussed in the following, determinism (of the usual physical variables \(A_i\) including the observed \(a_i\)) is a very strong assumption. However, the \textit{simultaneous definite values} (SDV) taken by two noncommuting observables is assumed by rather than proved by the EPR argument: the argument shows either the position or the momentum takes a definite value but it does \textit{not} show both simultaneously. Hence determinism in the EPR or Bell sense of SDV is \textit{not} inferred, contrary to the statement in [5, p.143]. It has to be assumed. But such an assumption violates QM which denies such possibility when the two corresponding quantum observables do not possess simultaneous eigenstate, which happens when they do not commute on their domains of definitions as in the case of the spin operators and the canonical position-momentum pair.

Determinism is also ruled out by various no-go theorems [11], [12] on hidden variables, which shows determinism is incompatible with the known predictions of quantum mechanics that we don’t doubt. So BT1 has a very weak starting point [4]. On the other hand, Bell argued [10] that (LC) follows directly from our intuitive physical concept of local causality (lc). Hence the argument of BT2 seems to start with a much broader assumption and does not need to bring in "locality" as in BT1 through the requirement of "separability" that no one ever doubts, apart from the "free will" issue [10, p. 244] which is not our concern.

Note that the EPR argument in (*) works differently in BT1 and BT2. In BT1 one cannot actually derive X=determinism from the EPR argument, as we indicated. On the other hand, (LC) is taken to be derivable from (lc), the latter is likely an assumption that EPR took to be obviously valid. We also take it to be valid in this paper, which may be
regarded as an alternative formulation of Einstein locality in which all interactions are local. In this paper we will analyze the hidden variable representation universally employed thus far in the Bell issue for general stochastic hidden variables. We will discuss the satisfaction of BI being equivalent to the existence of a joint probability distribution for the given spin variables, which has a long historical origin and bears strongly on the Bell issue through SDV. We will show the condition (LC) is \textit{not} a valid consequence of the physical conception (lc) except in the determinism limit. We will give a classical local correlation example that violates (LC) but not (lc), and produce a local hidden variable model that gives the correct quantum predictions for pair correlations from the singlet. We will show BT1 and BT2 are not equivalent on their faces, while the scope of validity of BT2 actually reduces to that of BT1. Since determinism violates QM in having SDV for noncommuting observables, or from the no-go theorems, BT1 and hence also BT2 have empty scope. Thus, the above argument from (\ast) is not valid for either Bell’s theorem and one can’t conclude the world or quantum physics is nonlocal in the sense of Bell nonlocality. Bell’s original derivation of the three-pair BI \cite{1} uses the fact that the spin values are either $+1$ or $-1$ ($+$ or $-1$). The inequality can be expressed more generally in terms of the probabilities of obtaining $++$ correlations in the pairs of spin as follows \cite[p. 220]{9},

\[ P_{13}(++) \leq P_{12}(++) + P_{23}(++) \]  

(1)

where the subscripts refer to the three spin directions. It is simpler to directly derive (1) from three of the four 3-variable joint distributions on the four spin values as given in Wigner \cite{13}. In fact, joint distribution derivation of BI was first given by Boole \cite{14}, \cite{15} and then a century later by Voboyev \cite{16} before Bell. Boole and Voboyev derived the necessary and sufficient conditions under which a joint distribution would exist among given 2-variable correlations on a finite set of random variables, including the cases of the 3- and 4-pairs BI. They cover also the results of Fine \cite{17}. Note that determinism corresponds to the degenerate limit of joint distribution.

The existence of a realizable joint distribution of noncommuting observables which gives the correct marginal single observable probabilities is itself a direct violation of QM since POVM measurements \cite{18} do not give correct marginal distribution on the single Hermitian observables. This contradiction with QM arises because nonzero probability is assigned in such a joint distribution to the elementary events that noncommuting observables take on simultaneous definite values, which by starting assumption are eigenvalues of the individual
observables. But as we noted there is no such simultaneous eigenstates in QM which a joint distribution would entail. Hence the model that leads to such joint distribution cannot be accepted as a legitimate possible hidden variable model. In such case, as in the squeezed state Wigner distribution representation of the original EPR two canonical pairs, BI would not be violated [19]. The Wigner distribution is positive but it is not the result of any measurement on any quantum state.

We next introduce the terminology of deterministic hidden variables (DHV) to denote the case in which each physical variable $\mathcal{A}_i$ is a function of some underlying hidden variable $\Lambda$ (with range in $\mathbb{R}^n$, the n-dimensional Euclidean space) with probability distribution $p(\lambda)$. (This DHV terminology varies in the literature. Also the term distribution here means general probability density including $\delta$-functions.) In [1] such an assumption is explicitly made. DHV differs little from determinism, as $\Lambda$ is the whole underlying variable but which we know just its distribution $p(\lambda)$. For DHV the probability of observing $\mathcal{A}$ is given by

$$P(a) = \int_{D_a} p(\lambda|\psi) \, d\lambda$$

(2)

where $P(a)$ is given by the Born rule for $\hat{A}$ measurement and state $\psi$, $D_a$ is the region of $\lambda$ with $\mathcal{A}(\lambda) = a$, and the distribution $p(\lambda)$ is transformed to the conditional distribution given $\psi$. A similar equation applies to the measurement of $\mathcal{B}$. In addition, we have the joint distribution $P(a,b)$ of obtaining values $a$ and $b$ simultaneous from joint measurement of variables $\mathcal{A}$ and $\mathcal{B}$ [20, p. 166],

$$P(a,b) = \int_{D_a \cap D_b} p(\lambda|\psi) \, d\lambda$$

(3)

where the region of integration is the intersection of $D_a$ and $D_b$. Note that $P(a,b)$ given by (3) is forced upon us by $\mathcal{A}(\lambda)$, $\mathcal{B}(\lambda)$, and $p(\lambda)$. What if the $\hat{A}$ and $\hat{B}$ corresponding to $\mathcal{A}$ and $\mathcal{B}$ do not commute? The DHV representation is simply inconsistent with quantum prediction in such case. Thus the DHV assumption has no scope of application as a hidden variable representation.

We have been using the principle that a hidden variable representation cannot lead to conclusion in contradiction with QM, which is implicit for the no-go theorems. On the other hand, the representation can be incomplete and does not yield all quantum predictions, say for POVM measurements, which is the case in all the ones that have been used by Bell and others. It is important to observe that SDV is the essential underlying feature in the hidden
variable issue. It is entailed by determinism, but determinism is not necessary for SDV since there may be a probability distribution of whatever origin on top. (See equ. (7) below.) Quantum mechanics embodies the principle of no SDV for noncommuting observables, not just the uncertainty relation.

The condition \( (LC) \) \[6, equ. (2)\] is, with a similar one switching \( a \) and \( b \),

\[
P(b|a, \lambda, \mathcal{C}) = P(b|\lambda, \mathcal{C})
\]  \hspace{1cm} (4)

The left side is the conditional probability of the B-site variable \( \mathcal{B} \) taking a measured value \( b \) given the A-site variable \( \mathcal{A} \) taking the value \( a \) conditioned on the hidden variable \( \Lambda \) taking value \( \lambda \), all with a common cause \( \mathcal{C} \) of the two spins. We do not need to pay attention to measurement settings of the two sites as they do not affect each other from Einstein locality. This \( (LC) \) condition is equivalent \[7\] to "factorizability",

\[
P(a, b|\lambda, \mathcal{C}) = P(a|\lambda, \mathcal{C}) P(b|\lambda, \mathcal{C})
\]  \hspace{1cm} (5)

We now show that physical local causality (lc) does not imply \( (LC) \), classically or quantum mechanically and for the same reason. It does only in the degenerate limit of determinism which violates quantum mechanics. First, consider the case of classical correlation in which a sequence of spin pairs is generated from a (local) source in which \( \sigma^A \) takes an angle \( \theta_A \) from 0 to \( 2\pi \) with any probability density, say the uniform density, on a fixed plane with a fixed axis. The spin \( \sigma^B \) is always opposite to \( \sigma^A \). Hence the joint density of the two spin directions is,

\[
P(\theta_A, \theta_B) = P(\theta_A) \delta(\theta_A - \theta_B - \pi)
\]  \hspace{1cm} (6)

It is evident that probability conditioning from (6) is not nonlocal action. Consider spin measurements on each of the two spins described by a measurement axis \( \mathcal{L} \) on the plane. If the spin direction points to the left of \( \mathcal{L} \) the result is \( + \), to the right the result is \( - \). (For the uniform density the probability is 0 that the spin points to \( \mathcal{L} \).) Thus perfect correlations are obtained between the two spins as noted in \[5\]. If we consider the case where \( P(\theta_B|\theta_A) \) is, for example, a continuously decreasing function of \( |\theta_B - \theta_A| \), then some correlation between the spin directions remain which is not perfect. See Fig. 1.
Figure 1: Orientation with respect to chosen measurement axis L. The result is $\pm$ depending on whether the spin is to the left or right of L as a vector.

The point is that in neither the perfect nor the imperfect correlation case is the probability (6) in the form (5), but both spins are locally correlated by the common source. The probability (6) would be in the form (5) only if $P(\theta_A)$ degenerates into deterministic $P(\theta_A) = \delta(\theta_A - \theta_0)$, or if the source actually produces a product density for $P(\theta_A, \theta_B|\mathcal{E})$, either of which is beyond local causality.

It is important to note that if we regard the random parameters $\theta_A$ and $\theta_B$ as hidden variables in a stochastic hidden variable (SHV) representation

$$P(a|\psi) = \int P(a|\lambda, \psi) p(\lambda)d\lambda$$

then one hidden variable, for example $\theta_A$, is still hidden in (7) when we take $\lambda = \theta_B$. The representation (7) used by Bell in [6] and his subsequent papers hides the random variable that may exist in specifying $a$ or its probability in addition to $\lambda$. This is in contrast with (2) in which $\lambda$ specifies all values of $\mathcal{A}(\lambda)$ in a DHV representation. There is a further generality in (7) compared to (2) in that the existence of a joint distribution (3) is not implied by (7) alone, thus (7) can be valid with no consequent contradiction with QM while (2) cannot be compatible with QM.

In the above example it is clear that everything is local. Thus it already shows clearly that (4) or (5) is not a consequence of the physical conception of local causality but is a lot more. Of course BI is satisfied in this classical situation, which is a consequence of the existence of joint distribution for the spin variable measurements. But whether (LC) follows from (lc) has nothing to do with whether the underlying physics is quantum or classical. The above situation also suggests that there could be a SHV representation for spin correlations in quantum mechanics which violates both BI and (LC). In the following we give such an example that gives correct quantum predictions.

The following local hidden variable model describes the singlet pair correlation measure-
ments as predicted by QM. We fix the measurement on site A to be the spin aligned with
the up state of the singlet. Generalization to arbitrary state and site A measurement for
tensor product measurements on the two qubits is given in the Appendix. Let \( \mathbf{m} \cdot \sigma^A \) and
\( \mathbf{n} \cdot \sigma^B \) be the spins of the A and B sites being measured with angle \( \theta \) between them. The
results \((+ -)\) and \((+++)\) are obtained when a uniformly distributed \( \Lambda \) takes value as follows,

\[
P(+, -) = (1 + \lambda_0)/2, \quad -\lambda_0 < \lambda < 1
\]

\[
P(+, +) = (1 - \lambda_0)/2, \quad 1 < \lambda < -\lambda_0
\]

where \( \lambda_0 = \cos^2 \theta \). This violates the BI (1) in accordance with QM. Of course (8)-(9) is a
SHV representation which does not satisfy (4) or (5), and it is local as \( \Lambda \) is a local random
parameter. Note that the randomness in the representation (7) with (8)-(9) is not reduced
to that of (2) as discussed above.

It may be stressed that the assumption (lc) underlying BT2 is more general than that
underlying BT1 which invokes DHV instead of SHV. Their equivalence by going through BI
[4] is not relevant since (LC) is not a consequence of (lc). Furthermore, forming product
distribution to get BI as in [17, equ. (12)] is incorrect as such product distribution does not
yield the correct given correlations.

In conclusion, we see that the mathematical local causality condition (LC) Bell used
does not follow from the physical conception of local causality except in the determinism
limit, while determinism can be ruled out a priori by the no-go theorems or simply by its
contradiction with the quantum principle of no simultaneous definite values, the latter also
underlies the GHZ paradox [18, p.152]. Hence, while Bell inequality violation as predicted
by quantum mechanics is a fact of nature, Bell’s theorem that quantum mechanics or nature
is nonlocal does not follow. In this connection, it may be noted (with detailed elaboration in
a future paper) that the EPR correlations and the "Einstein box" [9, §41] are already fully
explicated by probability conditioning via the Born rule without any need of the "projection
postulate" that seems to imply nonlocality. Probability conditioning is not action at a
distance.

What underlies all these is the just the fact that there are other forms of probability
theory from the usual ones, such as the one quantum mechanics provides. It would be
more fruitful to try delineate the newer form of probability that would yield the principle
of no simultaneous definite values, or provide an account in which the physical variables corresponding to quantum observables are emergent quantities, instead of insisting on a form of "realism" that quantum mechanics does not permit.

Acknowledgements

I would like to thank Derek Abbott for prompting me to look into Bell’s theorem.

Appendix

For a single qubit, the measurement probability in state $|\psi\rangle$ of a Hermitian spin observable $\hat{A} = \mathbf{m} \cdot \boldsymbol{\sigma}$, where $\mathbf{m}$ is a unit vector and $\boldsymbol{\sigma}$ a vector with Pauli operator components, is given by the representation (7) in which the hidden parameter $\Lambda$ is uniformly distributed,

\[
P(A = +) = \frac{1 + \lambda_0}{2} \quad \text{as } \lambda \text{ falls in the range } -\lambda_0 < \lambda < 1 \\
P(A = -) = \frac{1 - \lambda_0}{2} \quad \text{for } -1 < \lambda < -\lambda_0
\]  

(A1)  

(A2)

In (A1)-(A2) the constant $\lambda_0$ is given by $2p - 1$ where $p$ is the quantum mechanical probability that $a$ turns out $+$, i.e., $p = |\langle a = + | \psi \rangle|^2$ where $|a = +\rangle$ is the eigenstate of $\hat{A}$ with eigenvalue $+$. This representation is a simple generalization of the qubit hidden variable representation in average terms given in [1], [18, p. 159]. It covers all Hermitian qubit observables which can all be reduced to linear combination of the identity operator $I$ and the Pauli operators. For a general mixed state $\rho$ we can write in terms of the eigenstates states $|\psi_1\rangle$, $|\psi_2\rangle$, of $\rho$ and their probabilities, $p_1$, $p_2$,

\[
P(a) = p_1P(a|\psi_1\rangle + p_2P(a|\psi_2\rangle) 
\]  

(A3)

There would then be two hidden variables for the two eigenstates of $\rho$.

For two qubits in an arbitrary joint pure state we consider only general tensor product measurements $\mathbf{m} \cdot \boldsymbol{\sigma}^A \otimes \mathbf{n} \cdot \boldsymbol{\sigma}^B$ where $\mathbf{m}$, $\mathbf{n}$ are unit vectors and $\boldsymbol{\sigma}^A$, $\boldsymbol{\sigma}^B$ are Pauli operator vectors on the two sites. Let the A-site measurement result be $a$ with eigenstate $|a\rangle$, $\theta^B$ be the projection of $|\psi\rangle$ to $\mathcal{H}^B$ from $|a\rangle$. Then the joint probability for the $A=a$ and
A $B=b$ measurement for any given $|\psi\rangle$ on $\mathcal{H}^A \otimes \mathcal{H}^B$ can be specified by a hidden variable representation as follows: $P(A=a)$ given by (A1)-(A3), $P(B=b|A=a)$ given by (8)-(9) for the correlations which should yield the correct quantum prediction

$$P(B=b|A=a) = |\langle B = b|\theta^B\rangle|^2$$

(A4)

There are two hidden parameters (or a vector one) in this case, one effective for single qubit measurement and the other for correlations. In the case of the singlet state, we can choose $|\theta_i^A\rangle$ to align with $|\uparrow\rangle$. Thus (9) gives, for the $\hat{A}$ and $\hat{B}$ measurement which has angle $\theta$ between the $\mathbf{m}$ and $\mathbf{n}$ vectors of the two spins,

$$P(A=+, B=+) = \frac{1}{2} \sin^2 \theta = (1 - |\mathbf{m} \cdot \mathbf{n}|^2)/2$$

(A5)

Equation (A5) is the correct quantum result (9) that violates BI. The arbitrary joint state case goes by eigenstate decomposition as in the single qubit case.

Note that the above hidden variable model covers only tensor product spin measurements, which shares many features of the single qubit general Hermitian observable case and is not ruled out by no-go theorems. Nevertheless it shows clearly that local hidden variable needs not obey the (LC) condition although the physical (lc) condition is satisfied. Observe that all the hidden variable representation used in the Bell issue only give quantum results for some but not all measurements, say not the POVM ones. Thus our above example just shares this common incompleteness which we noted in the passage following equ. (3).

As a matter of fact, the simple representation (6) cannot be taken seriously as a hidden variable theory for several reasons, though it has long been so used as a schematic representation of hidden variables. As the above (A1)-(A2) show, such rewriting of the probability law from quantum mechanics by using quantum mechanics sheds no light on the underlying physics which is sought in a hidden variable theory. More significantly, the representation does not give a description of POVM measurements [18] which are contained within the formalism of quantum mechanics. Equally significantly, it may yield contradictions with QM, as in the case of giving joint distributions that do not exist as in (3). As we observe in the paper, such contradiction would make it an illegitimate hidden variable representation which no-go theorems or the principle of no SDV rule out as a matter of principle. In a future paper it will be shown that all the known no-go theorems can be derived from the quantum principle of no simultaneous definite values.
There are much more that can be said on what a meaningful, even if partial, hidden variable theory should be like. It appears this issue of what is a legitimate hidden variable approach is more significant than Bell inequality per se, and which actually first motivated Bell to look for physically meaningful conditions.

References


