On the Fatal Mistake Made by John S. Bell in the Proof of His Famous Theorem

Joy Christian*
Einstein Centre for Local-Realistic Physics, 15 Thackley End, Oxford OX2 6LB, United Kingdom
(Dated: 21 May 2016)

We explain the elementary mistake made by John S. Bell in the proof of his famous "theorem."

Consider – in slightly modern terms – the standard EPR-Bohm type experiment envisaged by John S. Bell to prove his famous theorem [1]. Alice is free to choose a detector direction $\mathbf{a}$ or $\mathbf{a}'$ and Bob is free to choose a detector direction $\mathbf{b}$ or $\mathbf{b}'$ to detect spins of the fermions they receive from a common source, at a space-like distance from each other. The objects of interest then are the bounds on the sum of possible averages put together in the manner of CHSH [2],

$$E(\mathbf{a},\,\mathbf{b})\,+\,E(\mathbf{a},\,\mathbf{b}')\,+\,E(\mathbf{a}',\,\mathbf{b})\,-\,E(\mathbf{a}',\,\mathbf{b}')\,, \tag{1}$$

with each average defined as

$$E(\mathbf{a},\,\mathbf{b})\,=\,\lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,A(\mathbf{a},\,\lambda^k)\;B(\mathbf{b},\,\lambda^k)\right]\,\equiv\,\Big\langle\,A_k(\mathbf{a})\,B_k(\mathbf{b})\,\Big\rangle\,, \tag{2}$$

where $A(\mathbf{a},\,\lambda^k)\equiv A_k(\mathbf{a})=\pm1$ and $B(\mathbf{b},\,\lambda^k)\equiv B_k(\mathbf{b})=\pm1$ are the measurement results of Alice and Bob, respectively. Now, since each $A_k(\mathbf{a})=\pm1$ and $B_k(\mathbf{b})=\pm1$, the average of their product is $-1\leqslant\Big\langle\,A_k(\mathbf{a})\,B_k(\mathbf{b})\,\Big\rangle\leqslant+1$. As a result, we can immediately read off the upper and lower bounds on the string of the four averages considered in equation (1):

$$-\,4\,\leqslant\,\Big\langle\,A_k(\mathbf{a})\,B_k(\mathbf{b})\,\Big\rangle\,+\,\Big\langle\,A_k(\mathbf{a})\,B_k(\mathbf{b}')\,\Big\rangle\,+\,\Big\langle\,A_k(\mathbf{a}')\,B_k(\mathbf{b})\,\Big\rangle\,-\,\Big\langle\,A_k(\mathbf{a}')\,B_k(\mathbf{b}')\,\Big\rangle\,\leqslant\,+\,4\,. \tag{3}$$

This should have been the final conclusion by Bell. But he continued. And in doing so he made one of his gravest mistakes. He replaced the above string of four separate averages of binary numbers with the following single average:

$$E(\mathbf{a},\,\mathbf{b})+E(\mathbf{a},\,\mathbf{b}')+E(\mathbf{a}',\,\mathbf{b})-E(\mathbf{a}',\,\mathbf{b}')\;\longrightarrow\;\Big\langle\,A_k(\mathbf{a})\,B_k(\mathbf{b})+A_k(\mathbf{a})\,B_k(\mathbf{b}')+A_k(\mathbf{a}')\,B_k(\mathbf{b})-A_k(\mathbf{a}')\,B_k(\mathbf{b}')\,\Big\rangle\,. \tag{4}$$

As innocuous as this step may seem, it is in fact an illegitimate step, for what is being averaged on the RHS are *un-observable*, and hence *un-physical* quantities. But this illegitimate step allows one to reduce the above average to

$$\Big\langle\,A_k(\mathbf{a})\,\{\,B_k(\mathbf{b})+B_k(\mathbf{b}')\,\}\,+\,A_k(\mathbf{a}')\,\{\,B_k(\mathbf{b})-B_k(\mathbf{b}')\,\}\,\Big\rangle\,. \tag{5}$$

And since each $B_k(\mathbf{b})=\pm1$, if $|B_k(\mathbf{b})+B_k(\mathbf{b}')|=2$, then $|B_k(\mathbf{b})-B_k(\mathbf{b}')|=0$, and vice versa. Consequently, using $A_k(\mathbf{a})=\pm1$, it is easy to conclude that the absolute value of the above average cannot exceed 2, just as Bell concluded:

$$-\,2\,\leqslant\,\Big\langle\,A_k(\mathbf{a})\,B_k(\mathbf{b})\,+\,A_k(\mathbf{a})\,B_k(\mathbf{b}')\,+\,A_k(\mathbf{a}')\,B_k(\mathbf{b})\,-\,A_k(\mathbf{a}')\,B_k(\mathbf{b}')\,\Big\rangle\,\leqslant\,+\,2\,. \tag{6}$$

Let us now try to understand why the replacement (4) is illegal. To begin with, Einstein's (or even Bell's own) notion of local-realism does not demand this replacement. The LHS of (4) satisfies the demand of local-realism perfectly well. To be sure, mathematically there is nothing wrong with a replacement of four separate averages with a single average. Every school child knows that the sum of averages is equal to the average of sums. But this rule of thumb is not valid in the above case, because $(\mathbf{a},\,\mathbf{b})$, $(\mathbf{a},\,\mathbf{b}')$, $(\mathbf{a}',\,\mathbf{b})$, and $(\mathbf{a}',\,\mathbf{b}')$ are *mutually exclusive pairs of measurement directions*, corresponding to four *incompatible* experiments. Each pair can be used by Alice and Bob for a given experiment, for all runs 1 to n, but no two of the four pairs can be used by them simultaneously. This is because Alice and Bob do not have the ability to make measurements along counterfactually possible pairs of directions such as $(\mathbf{a},\,\mathbf{b})$ and $(\mathbf{a},\,\mathbf{b}')$ simultaneously. Alice, for example, can make measurements along $\mathbf{a}$ or $\mathbf{a}'$, but not along $\mathbf{a}$ *and* $\mathbf{a}'$ at the same time.

But this inconvenient fact is rather devastating for Bell's argument, because it means that his replacement (4) above is illegitimate. Consider a specific run of the experiment and the corresponding quantity being averaged in (4):

$$A_k(\mathbf{a})\,B_k(\mathbf{b})\,+\,A_k(\mathbf{a})\,B_k(\mathbf{b}')\,+\,A_k(\mathbf{a}')\,B_k(\mathbf{b})\,-\,A_k(\mathbf{a}')\,B_k(\mathbf{b}')\,. \tag{7}$$

---

*Electronic address: `jjc@alum.bu.edu`

Here the index $k = 1$ now represents a specific run of the experiment. But since Alice and Bob have only two particles at their disposal for each run, only one of the four terms of the above sum is physically meaningful. In other words, the above quantity is physically meaningless, because Alice, for example, cannot align her detector along $\mathbf{a}$ and $\mathbf{a}'$ at the same time. And likewise, Bob cannot align his detector along $\mathbf{b}$ and $\mathbf{b}'$ at the same time. What is more, this would be true for all possible runs of the experiment, or equivalently for all possible pairs of particles. Which implies that all of the quantities listed below, as they appear in the average (5), are un-observable and hence physically meaningless:

$$
\begin{aligned}
&A_1(\mathbf{a})\,B_1(\mathbf{b}) \,+\, A_1(\mathbf{a})\,B_1(\mathbf{b}') \,+\, A_1(\mathbf{a}')\,B_1(\mathbf{b}) \,-\, A_1(\mathbf{a}')\,B_1(\mathbf{b}')\,,\\
&A_2(\mathbf{a})\,B_2(\mathbf{b}) \,+\, A_2(\mathbf{a})\,B_2(\mathbf{b}') \,+\, A_2(\mathbf{a}')\,B_2(\mathbf{b}) \,-\, A_2(\mathbf{a}')\,B_2(\mathbf{b}')\,,\\
&A_3(\mathbf{a})\,B_3(\mathbf{b}) \,+\, A_3(\mathbf{a})\,B_3(\mathbf{b}') \,+\, A_3(\mathbf{a}')\,B_3(\mathbf{b}) \,-\, A_3(\mathbf{a}')\,B_3(\mathbf{b}')\,,\\
&A_4(\mathbf{a})\,B_4(\mathbf{b}) \,+\, A_4(\mathbf{a})\,B_4(\mathbf{b}') \,+\, A_4(\mathbf{a}')\,B_4(\mathbf{b}) \,-\, A_4(\mathbf{a}')\,B_4(\mathbf{b}')\,,\\
&\quad\vdots\\
&A_n(\mathbf{a})\,B_n(\mathbf{b}) \,+\, A_n(\mathbf{a})\,B_n(\mathbf{b}') \,+\, A_n(\mathbf{a}')\,B_n(\mathbf{b}) \,-\, A_n(\mathbf{a}')\,B_n(\mathbf{b}')\,.
\end{aligned}
$$

But since each of the quantities above is physically meaningless, their average appearing on the RHS of (4), namely

$$\Big\langle\, A_k(\mathbf{a})\,B_k(\mathbf{b}) \,+\, A_k(\mathbf{a})\,B_k(\mathbf{b}') \,+\, A_k(\mathbf{a}')\,B_k(\mathbf{b}) \,-\, A_k(\mathbf{a}')\,B_k(\mathbf{b}')\,\Big\rangle\,, \tag{8}$$

is also physically meaningless. That is to say, no physical experiment can ever be performed — even by the God of Spinoza — that can meaningfully measure or evaluate the above average, since none of these quantities could have experimentally observable values. Therefore the innocuous looking replacement (4) made by Bell is in fact illegitimate.

On the other hand, it is important to note that each of the four averages appearing on the LHS of replacement (4),

$$E(\mathbf{a},\,\mathbf{b}) \,=\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,A(\mathbf{a},\,\lambda^k)\;B(\mathbf{b},\,\lambda^k)\right] \,\equiv\, \Big\langle\, A_k(\mathbf{a})\,B_k(\mathbf{b})\,\Big\rangle\,, \tag{9}$$

$$E(\mathbf{a},\,\mathbf{b}') \,=\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,A(\mathbf{a},\,\lambda^k)\;B(\mathbf{b}',\,\lambda^k)\right] \,\equiv\, \Big\langle\, A_k(\mathbf{a})\,B_k(\mathbf{b}')\,\Big\rangle\,, \tag{10}$$

$$E(\mathbf{a}',\,\mathbf{b}) \,=\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,A(\mathbf{a}',\,\lambda^k)\;B(\mathbf{b},\,\lambda^k)\right] \,\equiv\, \Big\langle\, A_k(\mathbf{a}')\,B_k(\mathbf{b})\,\Big\rangle\,, \tag{11}$$

$$\text{and}\quad E(\mathbf{a}',\,\mathbf{b}') \,=\, \lim_{n\,\gg\,1}\left[\frac{1}{n}\sum_{k\,=\,1}^{n}\,A(\mathbf{a}',\,\lambda^k)\;B(\mathbf{b}',\,\lambda^k)\right] \,\equiv\, \Big\langle\, A_k(\mathbf{a}')\,B_k(\mathbf{b}')\,\Big\rangle\,, \tag{12}$$

is a perfectly well defined and observable physical quantity. Therefore the bounds (3) on their sum are quite harmless. These bounds of $\{-4,\,+4\}$, however, have never been violated in any experiment (indeed, nothing can violate them).

In conclusion, Bell and his followers derive the upper bound of 2 on the CHSH string of averages by an illegal move. In the middle of their derivation they unjustifiably replace an observable, and hence physically meaningful quantity,

$$\Big\langle\, A_k(\mathbf{a})\,B_k(\mathbf{b})\,\Big\rangle \,+\, \Big\langle\, A_k(\mathbf{a})\,B_k(\mathbf{b}')\,\Big\rangle \,+\, \Big\langle\, A_k(\mathbf{a}')\,B_k(\mathbf{b})\,\Big\rangle \,-\, \Big\langle\, A_k(\mathbf{a}')\,B_k(\mathbf{b}')\,\Big\rangle\,, \tag{13}$$

with an experimentally un-observable, and hence physically entirely meaningless quantity

$$\Big\langle\, A_k(\mathbf{a})\,B_k(\mathbf{b}) \,+\, A_k(\mathbf{a})\,B_k(\mathbf{b}') \,+\, A_k(\mathbf{a}')\,B_k(\mathbf{b}) \,-\, A_k(\mathbf{a}')\,B_k(\mathbf{b}')\,\Big\rangle\,. \tag{14}$$

If they do not make this illegitimate replacement, then the upper bound on the CHSH string of averages is 4, not 2. It is mind-boggling why such a blatant mistake has been overlooked by the physics community for over fifty years [3].

[1] J. S. Bell, Physics **1**, 195 (1964); *Speakable and Unspeakable in Quantum Mechanics* (CUP, Cambridge, 1987), page 37.
[2] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[3] J. Christian, *Local Causality in a Friedmann-Robertson-Walker Spacetime*, `http://arXiv.org/abs/1405.2355` (2014).