Analysis of non locality proofs in Quantum Mechanics

Giuseppe Nisticò
Dipartimento di Matematica, Università della Calabria - Italy
and
INFN – gruppo collegato di Cosenza, Italy
E-mail: gnistico@unical.it

Abstract. Two kinds of non-locality theorems in Quantum Mechanics are taken into account: the theorems based on the criterion of reality and the quite different theorem proposed by Stapp. In the present work the analyses of the theorem due to Greenberger, Horne, Shimony and Zeilinger, based on the criterion of reality, and of Stapp’s argument are shown. The results of these analyses show that the alleged violations of locality cannot be considered definitive.

1. Introduction

The task accomplished by Quantum Mechanics as an empirical theory is to establish which relationships occur in Nature between physical events – including the occurrences of measurements’ outcomes – if the physical system is assigned a given state vector $|\psi\rangle$. All experimental observations so far performed have confirmed the quantum theoretical predictions which, per se, entail no violation of the locality principle we can express as follows.

(L) Principle of Locality. Let $R_1$ and $R_2$ be two space-time regions which are separated space-like. The reality in $R_2$ is unaffected by operations performed in $R_1$.

Conflicts between Quantum Mechanics and locality arise only if further conditions, which do not belong to the genuine set of quantum postulates, are required to hold.

In the “classical” non-locality theorems [1]-[3],[5]-[7] these further conditions bring back to the criterion of reality, introduced by Einstein, Podolsky and Rosen (EPR) [4]:

(R) Criterion of Reality. If, without in any way disturbing a system, we can predict with certainty the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.

EPR argued that under certain circumstances more non-commuting observables must have simultaneous physical reality as a consequence of (L) and (R) without being all measured, while Quantum Mechanics is unable to describe such a reality. This lack prompted to seek for a theory more complete than Standard Quantum Mechanics, able to ascribe reality to these unmeasured observables. But Bell [1] first, and other authors later [2][3], proved that such a local realistic theory cannot exist. More precisely, they found that contradictions arise just in the attempt of assigning values to non measured observables in agreement with (L), (R) and with Quantum Mechanics. Since the empirical validity of Quantum Mechanics cannot be denied,
these contradictions imply a violation of the condition \((L) \land (R)\), i.e. locality joined the criterion of reality.

A different approach leading to the need of faster than light transfer of information was followed by Stapp. He pointed out that the alleged demonstrations of Bell and followers above cited suffer a serious shortcoming: they “rest explicitly or implicitly on the local-hidden-variable assumption that the values of the pertinent observables exist whether they are measured or not. That assumption conflicts with the orthodox quantum philosophy” \[8\].

Then he developed \[9\] a non-locality proof which requires neither hidden variable hypothesis nor criteria of reality. According to this proof, locality is not consistent with the predictions of Quantum Mechanics about Hardy’s physical setting \[3\] if the following further assumptions are added to the standard quantum postulates.

- one assumption asserts that \textit{once a measurement outcome has actually occurred, no action in a space-like separated future region can change its value};
- the other assumption establishes that \textit{given a concrete specimen of the physical system, the choice of what observable to measure among the possible alternatives is free}.

The strategy pursued by this different non-locality theorem is to prove that the \textit{validity} of a specific statement \((SR)\), having the status of a physical law within Stapp’s approach, which concerns with the outcomes of measurements confined in a space-time region \(\mathcal{R}_\beta\), depends upon what it is freely chosen to do in a space-time region \(\mathcal{R}_\alpha\) separated space-like from \(\mathcal{R}_\beta\).

In the present work we show that the violation of locality, which results from the two kinds of theorems considered here, is not a definitive conclusion.

In the case of the theorems based on the criterion of reality \[1\]-\[3\], a recent analysis \[10\] highlighted that if the criterion of reality is interpreted according to its strict meaning, then their proofs fail. Instead, their proofs are valid if a \textit{wide} interpretation of the criterion of reality is assumed to hold. Therefore, these non-locality theorems can be interpreted as arguments against the wide interpretation and supporting the strict one, rather than as locality’s violations.

As stated in \[10\], the methods therein used for the analysis of the ‘classical’ non-locality theorems, which restores locality to Quantum Mechanics, become ineffective with respect to Stapp’s theorem because of the profound difference between the proof strategies. In this article we develop a methodology allowing for an analysis of Stapp’s proof on a logical ground. The results of our analysis show that a logical pitfall affects the proof; thus, the conclusion that locality is violated is not even reached by Stapp’s argument.

Section 2 of the present article introduces the basic theoretical concepts which enter the theorems at issue.

In section 3 we analyze the impact of the two different interpretations of the criterion of reality \((R)\) on the proof of a non-locality theorem based on such a criterion, namely the theorem of Greenberger, Horne, Shimony and Zeilinger \((\text{GHSZ})\) \[2\]. In so doing, first the different consequences of the strict and of the wide interpretation are identified in subsection 3.1. In subsection 3.2 it is shown that if the strict interpretation is assumed, then the proof of GHSZ is not successful.

In section 4 the logical structure of Stapp’s non-locality argument is described in detail. In section 5 we endow Stapp’s new assumptions \((\text{FC})\) and \((\text{NBITI})\) with formal content. This is necessary in order that a logico-mathematical analysis of the proofs of Property 1 and of Property 2, essential for the validity of the theorem, can be performed. The analysis of the proof of Property 1, shown in subsection 5.1, proves that it is correct. But since the proof of Property 2 turns out of be not valid, as we show in subsection 5.2, we conclude that according to the present analysis the non-locality theorem at issue is not valid.
The conclusive section 6 is devoted to relate the present work to other disproofs present in the literature.

2. Basic Formalism

Given a quantum state vector $|\psi\rangle$ of the Hilbert space $\mathcal{H}$ which describes the physical system, let $\mathcal{S}(|\psi\rangle)$ be a support of $\psi$, i.e. a concrete set of specimens of the physical systems whose quantum state is represented by $|\psi\rangle$. Let $A$ be any two-value observable, i.e. an observable having only two possible values denoted by $-1$ and $+1$, and hence represented by a self-adjoint operator $\hat{A}$ with purely discrete spectrum $\sigma(\hat{A}) = \{-1, +1\}$. Fixed any support $\mathcal{S}(|\psi\rangle)$, every two-value observable $\hat{A}$ identifies the following subsets $\mathcal{S}(|\psi\rangle)$:

- the set $A$ of the specimens in $\mathcal{S}(|\psi\rangle)$ which actually undergo a measurement of $A$;
- the set $A_+$ of the specimens of $A$ for which the outcome $+1$ of $A$ has been obtained;
- the set $A_-$ of the specimens of $A$ for which the outcome of $A$ is $-1$.

On the basis of the meaning of these concepts we can assume that the following statements hold (see [10], p.1268).

2. $A$ is a two-value observable then for all $|\psi\rangle$ a support $\mathcal{S}(|\psi\rangle)$ exists such that $A \neq \emptyset$;

2.ii $A_+ \cap A_- = \emptyset$ and $A_+ \cup A_- = A$;

2.iii $\langle \psi | \hat{A} | \psi \rangle \neq -1$ then $\mathcal{S}(|\psi\rangle)$ exists such that $A_+ \neq \emptyset$, and

\[
\langle \psi | \hat{A} \hat{\psi} \rangle \neq +1, \text{ then } \mathcal{S}(|\psi\rangle) \text{ exists such that } A_- \neq \emptyset;
\]

According to standard Quantum Theory, two observables $A$ and $B$ can be measured together if and only if $[\hat{A}, \hat{B}] = 0$; therefore also the following statements hold for every pair of two-value observables $A, B$.

2.iv $[\hat{A}, \hat{B}] \neq 0$ implies $A \cap B = \emptyset$ for all $\mathcal{S}(|\psi\rangle)$;

2.v $[\hat{A}, \hat{B}] = 0$ implies $\forall |\psi\rangle \exists \mathcal{S}(|\psi\rangle)$ such that $A \cap B \neq \emptyset$.

Given a pair $A, B$ of two-value observables such that $[\hat{A}, \hat{B}] = 0$, we say that the correlation $A \rightarrow B$ holds in the quantum state $|\psi\rangle$ if, whenever both $A$ and $B$ are actually measured, i.e. if $x \in A \cap B$, then $x \in A_+$ implies $x \in B_+$; so we have the following definition.

3. $A \rightarrow B$ if $[\hat{A}, \hat{B}] = 0$ and $x \in A_+$ implies $x \in B_+$, whenever $x \in A \cap B$.

This correlation admits the following characterization [11].

3.ii $A \rightarrow B \iff \frac{1 + \hat{A}}{2} \frac{1 + \hat{B}}{2} |\psi\rangle = \frac{1 + \hat{A}}{2} |\psi\rangle$.

Two observables $A$ and $B$ are separated, written $A \bowtie B$, if their respective measurements require operations confined in space-like separated regions $\mathcal{R}_\alpha$ and $\mathcal{R}_\beta$.

3. “Classical” non locality theorems

In this section we show how the non-locality theorems based on the criterion of reality fail if the criterion of reality (R) is interpreted according to its strict meaning. In so doing we limit ourselves to GHSZ theorem, because the “disproofs” [10] for the other theorems [1],[3] exploit the same ideas and methods.

In subsection 3.1 we deduce the implications which follow from the criterion of reality interpreted according to its strict meaning. Moreover, it is shown that stronger implications, like those required by the theorems of Bell and followers, can be deduced if a wide interpretation of the criterion is adopted.

In subsection 3.2 we show how the proof of GHSZ non-locality theorem cannot be successfully carried out if the strict interpretation of (R), instead of the wider one, is assumed to hold.
3.1. **Strict and Wide interpretation of EPR’s criterion**

We explain the two different interpretations of the criterion of reality by looking at the physical situation considered by EPR in [4], where they consider a system made up of two separated and non interacting sub-systems I and II. One of two non commuting observables $A$ and $B$ can be measured on system I, with non-degenerate eigenvalues $a_n$, $b_n$ and respective eigenvectors $\psi_n$, $\varphi_n$. Similarly, sub-system II possesses two non commuting observables $P$ and $Q$, with non-degenerate eigenvalues $p_k$, $q_k$ and respective eigenvectors $\upsilon_k$, $\nu_k$. The quantum state of the entire system I+II satisfies $\Psi = \sum_n \psi_n \otimes \upsilon_n = \sum_k \phi_k \otimes \nu_k$, so that, according to Quantum Mechanics, the following perfect correlations occur: if we actually measure $A$ (resp., $B$) on I obtaining the outcome $a_n$ (resp., $b_n$), then the outcome of an actual measurement of $P$ (resp., $Q$) on II is $p_n$ (resp., $q_n$). “Thus, by measuring either $A$ or $B$ we are in a position to predict with certainty, and without in any way disturbing the second system, either the value of the quantity $P$ [...] or the value of the quantity $Q$ [...].” Now, since $A$ and $B$ are non commuting, they cannot be measured together; therefore, the strict application of the criterion (R) leads to the following interpretation.

**Strict Interpretation.** Reality can be ascribed either to $P$ or to $Q$ according to which observable, either $A$ or $B$, is actually measured and whose outcome would allow for the prediction.

Instead, EPR’s attitude was different: “On the other hand, since at the time of measurement the two systems no longer interact [...] we arrived at the conclusion that two physical quantities $[P$ and $Q]$, with non-commuting operators, have simultaneous reality”. This means that in order to attain the simultaneous reality of $P$ and $Q$, EPR interpreted the criterion of reality as follows:

**Wide Interpretation.** For ascribing reality to $P$ (or $Q$) it is sufficient the “possibility” of performing the measurement of $A$ (or $B$) whose outcome would allow for the prediction, with certainty, of the outcome of a measurement of $P$ (or $Q$).

In order to express the two different interpretations within the theoretical apparatus, the formalism should be able to describe the reality, besides of the results of actually performed measurements, also of the “elements of reality” stemming from (R); hence, given $|\psi\rangle$ and fixed any support $\mathcal{S}(|\psi\rangle)$, we introduce the set $\tilde{A}$ of the descendents in $\mathcal{S}(|\psi\rangle)$ which objectively possess a value of the observable $A$, without being measured; by $\tilde{A}_+$ (resp., $\tilde{A}_-$) we denote the set of descendents of $\tilde{A}$ which possess the objective value +1 (resp., -1) of $A$; hence, we can assume that $\tilde{A}_+ \cap \tilde{A}_- = \emptyset$ and $\tilde{A}_+ \cup \tilde{A}_- = \tilde{A}$ hold. Then we define $A = \tilde{A} \cup A_+, \tilde{A}_+ = A_+ \cup A_+, \tilde{A}_- = A_- \cup A_-$. Of course, the “size” of $\tilde{A}$ depends, in general, on which interpretation of (R), the strict or the wide one, is adopted. Once defined the mappings $a: \tilde{A} \to \{1,-1\}$ and $\mathbf{a} : A \to \{1,-1\}$ by

$$ a(x) = \begin{cases} 1, & \text{if } x \in \tilde{A}_+ \\ -1, & \text{if } x \in \tilde{A}_- \end{cases} \quad \text{and} \quad \mathbf{a}(x) = \begin{cases} 1, & \text{if } x \in A_+ \\ -1, & \text{if } x \in A_- \end{cases}, $$

the correlation $A \to B$ can be equivalently expressed in terms of the mapping $\mathbf{a}$:

$$ A \to B \quad \text{if} \quad (a(x) + 1)(\mathbf{b}(x) - 1) = 0 \quad \text{for all } x \in A \cap B. $$

Now we can infer the implications of the strict interpretation of (R), we express as formal statements.

Let us suppose that $A \propto B$ holds, and that $A$ is measured on $x \in A$ obtaining $a(x) = 1$, i.e. $x \in \tilde{A}_+$. If the correlation $A \to B$ also holds, then the prediction of the outcome 1 can be considered valid for a measurement of $B$ on the same specimen. Now, by (L) the act of actually performing the measurement of $A$ does not affect the reality in $\mathcal{R}_3$; hence the criterion (R) could be applied to conclude that $x \in B$ and $b(x) = 1$:

$$ \text{if } A \propto B \text{ and } A \to B \text{ then } x \in A_+ \Rightarrow x \in B_+. $$
It is evident that this implication simply follows from the strict interpretation of the criterion (R); it can be more formally stated as follows.

(sR) if \( A \bowtie B \) and \( A \rightarrow B \) we can predict with certainty the value of an eventual measurement of \( B \) and ascribe reality to it once a measurement of \( A \) with concrete outcome \( a(x) = 1 \) is performed. If \( x \notin A_+ \) no prediction about \( B \) is allowed by (R) and (L).

Hence, according to (sR), \( A \bowtie B \) and \( A \rightarrow B \) imply \( A_+ \subseteq B_+ \subseteq B \) and the correlation \( (a(x) = 1) \Rightarrow (b(x) = 1) \), besides holding for all \( x \in A \cap B \), also holds for all \( x \in A_+ \). Analogously, if an actual measurement of \( B \) yields the outcome \(-1\), i.e. if \( x \in B_- \), then the strict interpretation of (R) leads us to infer that \( x \in A \) and \( a(x) = -1 \). Therefore it follows that \( B_- \subseteq A_- \subseteq A \) and that the correlation \( (a(x) = 1) \Rightarrow (b(x) = 1) \) also holds for every \( x \in B_- \). Hence, the correlation extends to \( A_+ \cup B_- \). Thus, from (R), (L) and Quantum Mechanics we infer the following statement.

(4.i) Extension of quantum correlations. Let \( A \) and \( B \) be space-like separated 2-value observables. If \( A \rightarrow B \) then

\[
(a(x) + 1)(b(x) - 1) = 0, \quad \forall x \in (A_+ \cup B_-) \cup (A \cap B).
\]

The quantum correlation \( A \leftrightarrow B \), i.e. \( A \rightarrow B \) and \( B \rightarrow A \), in the state \( \psi \) means that the correlation \( (a(x) = 1) \leftrightarrow (b(x) = 1) \) holds for all \( x \in A \cap B \) for all \( S(\psi) \). In this case, from (4.i) we can deduce that \( (a(x) = 1) \leftrightarrow (b(x) = 1) \) holds for all \( x \in (A_+ \cup B_-) \cup (B_+ \cup A_-) \cup (A \cap B) = A \cup B \), for all \( S(\psi) \). Hence, the strict interpretation of (R) also entails the following implications.

\[
A \bowtie B, \quad A \leftrightarrow B \quad \text{imply} \quad A \cup B \subseteq A \cap B \text{ i.e. } a(x) = b(x), \quad \forall x \in A \cup B, \quad \forall S(|\psi\rangle). \quad (4.ii)
\]

The wide interpretation of criterion (R) allows for larger extensions. Indeed it leads us to infer the following wider extensions of quantum correlations.

If \( A \bowtie B \) and \( A \rightarrow B \) then \( A_+ \subseteq B_+ \) and \( B_- \subseteq A_- \), \( \forall S(|\psi\rangle) \); \quad (5.i)

If \( A \bowtie B \) and \( A \leftrightarrow B \) then \( A_+ = B_+ \), \( B_- = A_- \) and \( A = S(|\psi\rangle) \), \( \forall S(|\psi\rangle) \). \quad (5.ii)

3.2. GHSZ theorem does not work with the strict interpretation

In this subsection we show how strong statements (5) implied by the wide interpretation play a decisive role in the non-locality theorem of GHSZ. But we show also that if we assume the strict interpretation, so that only the weaker statements (4) can be considered valid, then GHSZ proof fails.

GHSZ theorem makes use of seven two-value observables of a particular quantum system, divided into four classes

\[
\omega_A = \{A^0, A^1\}, \quad \omega_B = \{B\}, \quad \omega_C = \{C^0, C^1\}, \quad \omega_D = \{D^0, D^1\}.
\]

These observables have been singled out by GHSZ in such a way that

(6.i) two observables in two different classes commute and are separated from each other.

(6.ii) \( [\hat{A}^0, \hat{A}^1] \neq 0 \), \( [C^0, \hat{C}^1] \neq 0 \), \( [D^0, \hat{D}^1] \neq 0 \).

In general, provided that \( [\hat{A}, \hat{B}] = 0 \), by \( A \cdot B \) we denote the observable represented by the operator \( \hat{A}\hat{B} \); according to Quantum Theory, the product of the simultaneous outcomes of \( A \)
and $B$ is the outcome of $A \cdot B$. The state vectors $|\psi\rangle$ is chosen so that the following correlations between actually measured outcomes hold, according to Quantum Mechanics.

\begin{align*}
\text{i)} & \quad a^\alpha(x) b(x) = -c^\alpha(x) d^\alpha(x) \quad \forall x \in (A^\alpha \cap B) \cap (C^\alpha \cap D^\alpha) \equiv X, \\
\text{ii)} & \quad a^\beta(y) b(y) = -c^\beta(y) d^\beta(y) \quad \forall y \in (A^\beta \cap B) \cap (C^\beta \cap D^\beta) \equiv Y, \\
\text{iii)} & \quad a^\beta(z) b(z) = -c^\alpha(z) d^\beta(z) \quad \forall z \in (A^\beta \cap B) \cap (C^\alpha \cap D^\beta) \equiv Z, \\
\text{iv)} & \quad a^\alpha(t) b(t) = c^\beta(t) d^\beta(t) \quad \forall t \in (A^\alpha \cap B) \cap (C^\beta \cap D^\beta) \equiv T. 
\end{align*}

Equations (7.i), (7.ii), (7.iii), (7.iv) express the perfect quantum correlations $A^\alpha \cdot B \mapsto -C^\alpha \cdot D^\alpha$, $A^\beta \cdot B \mapsto -C^\beta \cdot D^\beta$, $A^\alpha \cdot B \mapsto -C^\alpha \cdot D^\beta$, $A^\beta \cdot B \mapsto C^\beta \cdot D^\alpha$, respectively.

According to the wide interpretation, (5.ii) holds and therefore correlations (7) can be extended to the following correlations between objective values.

\begin{align*}
\text{i)} & \quad a^\alpha(x) b(x) = -c^\alpha(x) d^\alpha(x), \\
\text{ii)} & \quad a^\beta(y) b(y) = -c^\beta(y) d^\beta(y), \\
\text{iii)} & \quad a^\beta(z) b(z) = -c^\alpha(z) d^\beta(z), \\
\text{iv)} & \quad a^\alpha(t) b(t) = c^\beta(t) d^\beta(t) \quad \forall x \in S(|\psi\rangle). 
\end{align*}

The contradiction proved by GHSZ lies just in (8). Indeed, given any $x \in S(|\psi\rangle) \neq \emptyset$, from (8.i) and (8.iv) we get

$$c^\alpha(x) d^\alpha(x) = -c^\beta(x) d^\beta(x).$$

From (8.iii) and (8.ii) the equality $c^\alpha(x) d^\beta(x) = c^\beta(x) d^\alpha(x)$ follows, which is equivalent to

$$c^\alpha(x) d^\alpha(x) = c^\beta(x) d^\beta(x)$$

which contradicts (9).

Now we prove that this GHSZ proof of inconsistency does not work if we replace the implications (5) by the weaker (4) allowed by the strict interpretation. The extension of correlations (7) implied by (4.ii) is the following.

\begin{align*}
\text{i)} & \quad a^\alpha(x) b(x) = -c^\alpha(x) d^\alpha(x) \quad \forall x \in (A^\alpha \cap B) \cup (C^\alpha \cap D^\alpha) \equiv \bar{X}, \\
\text{ii)} & \quad a^\beta(y) b(y) = -c^\beta(y) d^\beta(y) \quad \forall y \in (A^\beta \cap B) \cup (C^\beta \cap D^\beta) \equiv \bar{Y}, \\
\text{iii)} & \quad a^\beta(z) b(z) = -c^\alpha(z) d^\beta(z) \quad \forall z \in (A^\beta \cap B) \cup (C^\alpha \cap D^\beta) \equiv \bar{Z}, \\
\text{iv)} & \quad a^\alpha(t) b(t) = c^\beta(t) d^\beta(t) \quad \forall t \in (A^\alpha \cap B) \cup (C^\beta \cap D^\beta) \equiv \bar{T}. 
\end{align*}

In order that the GHSZ argument – which leads to the contradiction from (8) to (10) through (9) – can be successfully repeated starting from (11), the first step requires that (11.i) and (11.iv) should hold for the same specimen $x_0$; therefore the condition $\bar{X} \cap \bar{T} \neq \emptyset$ should hold; the second step requires that also (11.ii) and (11.iii) should hold for such a specimen $x_0$. Thus, the condition

$$\bar{X} \cap \bar{Y} \cap \bar{Z} \cap \bar{T} \neq \emptyset$$

should be satisfied. Now, from (6.ii) and (2.ii) we derive

$$\emptyset = (A^\alpha \cap B) \cap (A^\beta \cap B) = (C^\alpha \cap D^\alpha) \cap (C^\beta \cap D^\alpha) = (13)$$

$$= (C^\alpha \cap D^\alpha) \cap (C^\alpha \cap D^\beta) = (C^\alpha \cap D^\beta) \cap (C^\beta \cap D^\beta) = (C^\beta \cap D^\beta) \cap (C^\beta \cap D^\beta).$$

By making use of (11) and (13) we deduce $\bar{X} \cap \bar{Y} \cap \bar{Z} \cap \bar{T} = \emptyset$, which refutes condition (12) necessary to prove the inconsistency. Thus, GHSZ proof fail if the strict interpretation replaces the wide one.
4. A different non locality theorem
In this section we formulate in detail the argument proposed by Stapp to show that Quantum mechanics violates locality without making use of hidden variable hypotheses or criteria of reality.

Let us first establish the three hypotheses of Stapp’s theorem

(FC) Free Choices: “This premise asserts that the choice made in each region as to which experiment will be performed in that region can be treated as a localized free variable.”[9]

(NBITI) No backward in time influence: “This premise asserts that experimental outcomes that have already occurred in an earlier region [...] can be considered fixed and settled independently of which experiment will be chosen and performed later in a region spacelike separated from the first.”[9]

The third premise of Stapp’s theorem affirms the existence, as established by Hardy[3], of four two-value observables $A^{(1)}$, $A^{(2)}$, $B^{(1)}$, $B^{(2)}$ and of a particular state vector $|\psi\rangle$ for a certain physical system, which satisfy the following conditions:

(h.i) $A^{(1)}$, $A^{(2)}$ are confined in a region $R_{\alpha}$ separated space-like from the region $R_{\beta}$ wherein the observables $B^{(1)}$ and $B^{(2)}$ are confined, with $R_{\alpha}$ lying in time earlier than $R_{\beta}$. Hence in particular $A^{(j)} \bowtie B^{(k)}$, $j, k \in \{1, 2\}$.

(h.ii) $[\hat{A}^{(1)}, \hat{A}^{(2)}] \neq 0$, $[\hat{B}^{(1)}, \hat{B}^{(2)}] \neq 0$; $-1 \neq \langle \psi | \hat{A}^{(j)} | \psi \rangle \neq +1$, $-1 \neq \langle \psi | \hat{B}^{(j)} | \psi \rangle \neq +1$.

(h.iii) $[\hat{A}^{(j)}, \hat{B}^{(k)}] = 0$, $j, k \in \{1, 2\}$, and in the state vector $|\psi\rangle$ the following chain of correlations holds.

\begin{align*}
& a) A^{(1)} \rightarrow B^{(1)}, \quad b) B^{(1)} \rightarrow A^{(2)}, \quad c) A^{(2)} \rightarrow B^{(2)}. \\
& (h.iv) S(|\psi\rangle) \text{ and } x_0 \in S(|\psi\rangle) \text{ exist such that } x_0 \in A_{+}^{(1)} \cap B_{-}^{(2)}. 
\end{align*}

In fact, this last condition is implied from the following non-equality satisfied by Hardy’s setting.

$$\left\langle \psi \mid \frac{1}{2} \hat{A}^{(1)} - \frac{1}{2} \hat{B}^{(2)} \psi \right\rangle \neq 0. \quad (14)$$

Since the l.h.s. is nothing else but the quantum probability that a simultaneous measurement of $A^{(1)}$ and $B^{(2)}$ yields respective outcomes $+1$ and $-1$, the non-equality states that the correlation $A^{(1)} \rightarrow B^{(2)}$ does not hold. Therefore, by (3.i) it implies (h.iv).

The logical mechanism of the non-locality proof at issue is based on the following pivotal statement.

(SR) “If $[B^{(1)}]$ is performed and gives outcome $[+1]$, then, if, instead, $[B^{(2)}]$ had been performed the outcome would have been $[+1]$.”[9]

By leaving out for the time being the question of its validity, we have to recognize, following Stapp, that (SR) has the status of a physical law about outcomes of measurements completely performable within region $R_{\beta}$. Then Stapp introduces the following statements.

Property 1. If a measurement of $A^{(2)}$ is performed in region $R_{\alpha}$, then (SR) is valid.

In formula,

$$x \in A^{(2)} \Rightarrow (SR) \text{ holds for this } x.$$

Property 2. If a measurement of $A^{(1)}$ is performed in region $R_{\alpha}$, then (SR) is not valid.

In formula,

$$x \in A^{(1)} \not\Rightarrow (SR) \text{ holds for this } x.$$
If both these properties actually followed from the premises (FC), (NBITI), (h.i-iv), then the validity of statement (SR) would depend on what is decided to do in region $R_\alpha$, separated space-like from $R_\beta$; hence a violation of the following locality principle would happen:

“The free choice made in one region as to which measurement will be performed there has, within the theory, no influence in a second region that is spacelike separated from the first.” [9]

In fact, Stapp gives his own proofs [9] that both property 1 and property 2 do hold. Thus, we should conclude that the above locality principle is violated if the three premises hold.

5. Logical analysis

In this section we shall examine, from a mere logical point of view, the proofs of property 1 and property 2 as drawn by Stapp. Let us begin by considering property 1.

Property 1: $x \in A^{(2)}$ implies (SR) holds for this $x$.

Stapp’s Proof: “The concept of ‘instead’ [in (SR)] is given a unambiguous meaning by the combination of the premises of ‘free’ choice and ‘no backward in time influence’: the choice between $[B^{(2)}]$ and $[B^{(1)}]$ is to be treated, within the theory, as a free variable, and switching between $[B^{(2)}]$ and $[B^{(1)}]$ is required to leave any outcome in the earlier region $[R_\alpha]$ undisturbed. But the statements [(h.iii.a) and (h.iii.b)] can be joined in tandem to give the result (SR)” [9].

We see that the steps of this proof are carried out by appealing to their intuitiveness, rather than by means of the usual logico-mathematical methods, so that in this form the proof unfits for an analysis on a logical ground. In particular, the possibility of such an analysis would require that the “unambiguous meaning of the concept of ‘instead’ ” be endowed with a mathematical counterpart within the theoretical apparatus, in order to make explicit its role and formally verifiable the proof.

We provide such a mathematical counterpart by means of a precise implication which can be inferred from the premises (FC) and (NBITI) for two separated observables $A$ and $B$, respectively confined in space-like separated regions $R_\alpha$ and $R_\beta$, with $R_\alpha$ lying in time earlier than $R_\beta$, such that the empirical implication $A \rightarrow B$ holds in the state $|\psi\rangle$.

Given any concrete specimen $x \in S(|\psi\rangle)$, the validity of condition (FC) makes sensible the question:

“what would be the outcome of a measurement of $B$?”

also in the case that $B$ is not measured on that particular specimen $x$, independently of which, if any, observable is measured in region $R_\alpha$. This meaningfulness forces the introduction of two further extensions $B_+$ and $B_-$ in $S(|\psi\rangle)$ of any two-value observable $B$ confined in $R_\beta$.

The extension $B_+$ (resp. $B_-$) is defined to be the set of the specimens $x \in S(|\psi\rangle)$ such that if $B$ had been measured, even instead of an actually measured observable $C$ in $R_\beta$, then outcome +1 (resp. –1) would have occurred.

In general, a prediction of which specimens belong to $B_+$ or to $B_-$ is not possible, but the coherence of the new concepts requires that the following statement hold.

\[(15.i) \quad B_+ \cap B_- = \emptyset;\]

\[(15.ii) \quad x \in B_- \Rightarrow x \notin B_+ \quad \text{and} \quad x \in B_+ \Rightarrow x \notin B_-\]

Now we make use of (NBITI), by taking into account that the correlation $A \rightarrow B$ holds. If $A$ is actually measured on $x \in S(|\psi\rangle)$ and the outcome +1 is obtained, i.e. if $x \in A_+$, such a value does not depend, because of (NBITI), on the choice of what is decided to measure in $R_\beta$. Since
A → B we have to conclude that if B were measured on that specimen x then the outcome +1 would be obtained. Thus, we have inferred the following implication from the premises (FC) and (NB)TI.

\[(15.iii) \quad \text{If } A \preceq B \text{ and } A \rightarrow B \text{ then } x \in A_+ \Rightarrow x \in B_+.
\]

The new theoretical concepts just introduced make possible to re-formulate the crucial statement (SR) of Stapp’s argument in the following very simple form.

\[(SR) \quad x \in B_+^{(1)} \implies x \in B_+^{(2)}.
\]

5.1. Property 1.

Now we can analyze the proof of Property 1, by expanding it in the following sequence of statements.

\[(E.1) \quad \text{Let us suppose that the antecedent of Property 1 holds:} \\
\quad x \in A^{(2)}.
\]

\[(E.2) \quad \text{Let us suppose that the antecendent of (SR) holds too:} \\
\quad x \in B_+^{(1)}.
\]

\[(E.3) \quad \text{Hence (16.i) and (16.ii) imply} \\
\quad x \in B^{(1)} \cap A^{(2)}.
\]

\[(E.4) \quad \text{Then (h.iii), (16.ii) and (16.iii) imply} \\
\quad x \in A_+^{(2)}.
\]

\[(E.5) \quad (\text{h.iii.c), (16.iv) and (15.iii) imply} \\
\quad x \in B_+^{(2)}.
\]

In order that this re-worded proof be correct, it is sufficient to prove that specimens satisfying (16.i) and (16.ii) actually exist, since the steps from (E.3) to (E.5) are correctly demonstrated. Now, by (h.iii.b), (3.ii) and (h.ii) we have \(\frac{1+B^{(1)}}{2} \frac{1+A^{(2)}}{2} \psi = \frac{1+B^{(1)}}{2} \psi \neq 0\). Therefore \(\langle \psi | \frac{1+B^{(1)}}{2} \frac{1+A^{(2)}}{2} \psi \rangle \neq 0\). But this last is just the probability that a simultaneous measurement of \(B^{(1)}\) and \(A^{(2)}\) yields respective outcomes +1 and +1; being it non vanishing we have to conclude that a specimen \(x\) satisfying (16.i) and (16.ii) actually exists.

Thus our analysis does agree with Stapp’s conclusion that (SR) holds if \(A^{(2)}\) is measured in \(\mathcal{R}_\alpha\).

5.2. Property 2.

Now we submit the proof of property 2 to our analysis.

Property 2: \(x \in A^{(1)}\) does not imply \((SR)\) holds for this \(x\).

Hence this time Stapp’s scope is to show that \(x_0 \in A^{(1)}\) exists such that the antecedent of \((SR)\) is true but the consequent is false, i.e. that

\[\exists x_0 \in A^{(1)}, \quad x_0 \in B_+^{(1)} \text{ but } x_0 \notin B_+^{(2)}.
\]

(17)
Stapp’s Proof: “Quantum theory predicts that if \([A^{(1)}]\) is performed, then outcome \([+1]\) appears about half the time. Thus, if \([A^{(1)}]\) is chosen, then there are cases where \([x \in A^{(1)}_+]\) is true. But in a case where \([x \in A^{(1)}_+]\) is true, the prediction [(h.iii.a)] asserts that the premise of (SR) is true. But statement [(h.iv)], in conjunction with our two premises that give meaning to ‘instead’, implies that the conclusion of (SR) is not true: if \([B^{(2)}]\) is performed instead of \([B^{(1)}]\), the outcome is not necessarily \([+1]\), as it was in case \([A^{(2)}]\). So, there are cases where \([A^{(1)}]\) is true but (SR) is false.” [9]

Conclusion (17) is attained by Stapp through the following sequence of statements we translate from his proof.

(S.1) A support \(S(\psi)\) exists such that \(A^{(1)}_+ \neq \emptyset\).
(S.2) \(x \in A^{(1)}_+ \Rightarrow x \in B^{(1)}_+\).
(S.3) The antecedent of (SR) holds \(\forall x \in A^{(1)}_+\).
(S.4) \(\exists x_0 \in A^{(1)}_+\) such that \(x_0 \in B^{(2)}_-\).
(S.5) \(x_0 \notin B^{(2)}_-\).

Let us now check the validity of each step.

Statement (S.1) holds by (2.iii) and (h.ii).
Statement (S.3) is implied from (S.1) and (S.2).
Statement (S.4) holds because of (h.iv).
Statement (S.5) holds because of (S.4) and (15.ii).

We see that all steps (S.1), (S.3), (S.4), (S.5) hold true according to a logical analysis.

What about step (S.2)? Statement (S.2) is nothing else but the translation into our language of the phrase “But in a case where \([x \in A^{(1)}_+]\) is true, the prediction [(h.iii.a)] asserts that the premise of (SR) is true” stated by Stapp in his proof. Hence, according to Stapp’s proof, (S.2) holds because of (h.iii.a) \(A^{(1)} \rightarrow B^{(1)}\). But the implication

\[x \in A^{(1)}_+ \Rightarrow x \in B^{(1)}_+\]

follows from \(A^{(1)} \rightarrow B^{(1)}\) if \(x \in A^{(1)}_+ \cap B^{(1)}_+\) holds too, because of (3.i). However, this last condition cannot hold for the specimen \(x_0\) considered in (S4), because it has been characterized by the two conditions \(x_0 \in A^{(1)}_+\) and \(x_0 \in B^{(2)}_-\). But if \(x_0 \in B^{(2)}_-\) holds then \(x_0 \in B^{(2)}\) obviously holds too, so that the premise of (SR) \(x_0 \in B^{(1)}_+\) cannot hold because \(B^{(1)}\) and \(B^{(2)}\) do not commute with each other and therefore \(B^{(1)} \cap B^{(2)} = \emptyset\), by (h.ii) and (2.iv).

6. Conclusive remarks

In this work we have analyzed two kinds of theorems proposed in the literature for proving that the principle of locality is not consistent with Quantum Mechanics. Since Quantum Mechanics per se, i.e. without adding further assumptions to the genuine quantum postulates, does not conflict with locality, every non-locality theorem can reach the aimed inconsistency only by introducing some other conditions besides the standard ones.

In the first kind of non-locality theorems, like the theorem of Bell [1], these further conditions can be identified with the criterion of reality established by EPR in their famous 1935 paper.
[4]. Now, in [10] it has been put forward that the interpretation of this criterion is not unique. As shown in section 3, the interpretation of EPR goes beyond the strict meaning of the criterion. The non-locality theorems assuming the criterion of reality are successful if this wide interpretation is adopted. But we show in section 3.2 that if the criterion is interpreted according to its strict meaning then the non-locality proof of GHSZ [2] becomes unable to reach the inconsistency. Similar disproofs for the other non-locality theorems based on the criterion of reality can be found in [10].

The argument proposed by Stapp aims to prove inconsistency between Quantum Mechanics and locality by avoiding the use of criteria of reality or hidden variable hypotheses, because they entail, contrary to quantum philosophy, the assignment of pre-existing values to observables which are not measured. In the present work we have analyzed the final version of Stapp’s proof, published in [9], the author recognizes as the more effective. In fact, such a final form is the result of a number of works started in 1975 [12], submitted to various improvements over the years. These works received severe criticisms [13]-[16], all answered by Stapp [17]-[19].

However, the debate has not reached a definitive conclusion, because the criticisms enter the counterfactual character of the concept of “instead” used in Stapp’s argument and their aim is to check the validity of the proof within counterfactuals theory, i.e. modal logic [20]. On the other hand, in his replies Stapp maintains that his proof, contrary to the earliest versions, does not make use of modal logic.

The analysis presented in the present works does not make use of counterfactuals theory. Indeed, our disproof proceeds

- first, by translating the consequences of Stapp’s further assumptions (FC) and (NB1TI) into the formal statements (15.1)-(15.iii) within a suitable theoretical apparatus able to describe Stapp’s approach. No counterfactual concepts, such as “possible worlds” or “nearness of possible worlds”, are involved in such a translation;
- then, the proofs of property 1 and property 2 as drawn by Stapp are analyzed from an ordinary (not modal) logico-mathematical point of view.

Since the proof of property 2, at the end of the analysis, turns out to be not valid, we have to conclude that Stapp’s argument fails within our theoretical apparatus.

Thus, Stapp’s refusals of previous criticisms do not apply to the disproof presented in the present work.

References
[1] Bell J S 1964 Physics 1 165
[2] Greenberger D M, Horne M A, Shimony A and Zeilinger A 1990 Am. J. Phys. 58 1131
[3] Hardy L 1993 Phys. Rev. Lett. 71 1665
[4] Einstein A, Podolsky B and Rosen N 1935 Phys. Rev. 47 777
[5] Clauser J F, Horne M A, Shimony A and Holt R A 1969 Phys. Rev. Lett. 23 880
[6] Mermin N D 1993 Rev. Mod. Phys. 65 803
[7] Mermin N D 1995 Phys. Rev. Lett. 74 831
[8] Stapp H P 2006 Found. Phys. 36 73
[9] Stapp H P 2004 Am. J. Phys. 72 30
[10] Nisticò G and Sestito A 2011 Found. Phys. 41 1263
[11] Nisticò G 1995 Found. Phys. 25 1757
[12] Stapp H P 1975 Nuovo Cimento 29 270
[13] Shimony A 2006 Found. Phys. 36 61
[14] Clifton R K and Dickson M 1994 Phys. Rev. A 49 4251
[15] Shimony A and Stein H 2001 Am. J. Phys. 69 848
[16] Mermin N D 1998 Am. J. Phys. 66 920
[17] Stapp H P 1994 Phys. Rev. A 49 4257
[18] Stapp H P 1998 Am. J. Phys. 66 924
[19] Stapp H P 2006 *Found.Phys.* **36** 73
[20] Lewis D 1973 *Counterfactuals* (Cambridge, MA: Harvard University Press)