Review of Nelson’s analysis of Bell’s theorem

Benjamin Schulz, *
February 19, 2013

Abstract

This article contains a review of Nelson’s analysis of Bell’s theorem. It shows that Bell’s inequalities can be violated with a theory of local random variables if one accepts that the outcomes of these variables are not predetermined prior to measurement. Furthermore, there can be no model of the singlett state at all where the outcomes do not explicitly depend on the settings of the measurement devices. The article describes the relations between Bell’s theorem and the Free Will theorem of Conway and Kochen. The article closes by relating the various assumptions needed to derive Bell’s theorem with the reality criterion of EPR.

Contents

1 Introduction .................................................. 1
2 Nelson’s analysis of Bell’s theorem ...................... 2
   2.1 The setup of the EPR experiment in theory .......... 2
   2.2 Active locality .......................................... 4
   2.3 Passive locality ........................................ 5
   2.4 Nelson’s theorem ....................................... 5
3 Implications of passive locality .......................... 8
4 The relation between Bell’s theorem and the Free Will theorem of Conway and Kochen .......... 9
5 Conclusions .................................................. 12
6 Relation between the assumptions behind Bell’s inequality and the locality principle of Einstein ........ 12

1 Introduction

In 1935, Einstein, Podolsky and Rosen (EPR) wrote an article in which they denied that quantum theory would be a complete theory of nature [1]. Around 1951, Bohm gave a more testable outline of the so-called “EPR paradox” [2, 3]. He described a thought experiment with one source that ejects particles having opposite spin to two spatially separate Stern-Gerlach magnets of variable orientations, see Fig. 1. Then, in 1964, Bell published a theorem about this paradox in the form of an inequality. It made clear that hidden variable theories fulfilling certain conditions would contradict quantum mechanics. Bell called these conditions “locally causal” and explained them in detail in [4].

In his first publication on that topic, Bell wrote: “If hidden parameters would be added to quantum mechanics, there must be a mechanism, whereby the setting of one measuring device

*Connollystr. 3, 80809 München, e-mail: Benjamin.Schulz@physik.uni-muenchen.de
can influence another spatially separated device” and the signal involved has to “propagate instantaneously” [5]. However, Bell’s first contribution underwent several modifications. Over the years, more and more instructive proofs of his inequality were constructed by him and others. A collection of all of Bell’s fundamental articles can be found in [7]. Finally in 1969, Clauser, Holt, Shimony and Horne (CHSH) brought the inequality of Bell into a form suitable for experimental investigation [8].

In 1985, Nelson tried to analyse Bell’s theorem with full mathematical rigour. The result of his analysis was that Bell’s definition of “locally causal”, which is the starting point to derive Bell’s inequality, could be divided into two separate conditions. Both were necessary to derive the inequality, but only one of them has to hold if in a hidden variable theory an outcome depends on changes made at a spatially separated location. Nelson published his result twice [9, 10]. (However, a small correction was added later [11], in order to make Nelson’s theorem compatible with Mermin’s presentation [12].) Furthermore, there are studies from other mathematicians, for example [13], which give further insightful analysis of Nelson’s work. Unfortunately, Nelson’s articles on Bell’s inequality got almost overlooked by physicists. The reason for this might be that Nelson found his own interpretation of quantum mechanics [14, 15] to be at variance with the requirements of his theorem for a model without instantaneous signalling effects.

This article is organised as follows: In section 2, we review Nelson’s contribution towards a mathematically rigorous understanding of Bell’s inequality. The article will need some understanding of mathematical probability theory. For a general introduction to this theory, see [16, 17, 18, 19].

2 Nelson’s analysis of Bell’s theorem

2.1 The setup of the EPR experiment in theory

An EPR experiment consists of two measurement devices 1 and 2 which are space-like separated, as well as a particle source in the intersection of their past cones. The source ejects pairs of particles to the two detectors. At 1 and 2, the direction of the spins is measured with a Stern-Gerlach magnet. These magnets can be rotated around arbitrary directions. The direction of the magnets is called $\vec{\mu}$ for detector 1 and $\vec{\nu}$ for detector 2 (see Figs. 1, 2).

When the particles arrive in the Stern-Gerlach magnets, the magnetic field of these devices could change the particle’s properties, including the spin. In an EPR experiment, the parts of the detectors that might influence the particles are the axes $\vec{\mu}$ and $\vec{\nu}$. Those axes can be chosen by the experimenter freely at will. Due to the Stern-Gerlach magnets, we have to describe our measurement results by device dependent random variables. For the two detectors 1 and 2, we define two axis dependent families of random variables $D_{31\vec{\mu}} : \Omega_1 \rightarrow \Omega_{1D}$ and $D_{32\vec{\nu}} : \Omega_2 \rightarrow \Omega_{2D}$.

The outcomes of these random variables are $\Omega_{1D} = \Omega_{2D} = \{\uparrow, \downarrow\}$. Since the random variables
$D_{1\beta}$ and $D_{2\alpha}$ depend on the setting of the Stern-Gerlach magnets, we write their outcomes on the state space using the axis dependent notation $\tilde{F} \equiv P(\Omega)$. The probability measure of the enlarged probability space will be denoted by $P$.

The results at the detectors could depend on the preparation of the particles by the source. We denote the set of outcomes at the source as $\Omega_S$ and say that the corresponding events happen at preparation stage. Since we have a set of outcomes at the source and at the detectors, we must describe the EPR experiment with an enlarged probability space. We denote the outcomes of that space as $\Omega = \Omega_D \times \Omega_S \times \Omega_{2D}$ and define a sigma algebra $\mathcal{F}$ on that set by the power set $\mathcal{F} \equiv P(\Omega)$. The probability measure of the enlarged probability space will be denoted by $P$.

On this enlarged probability space, we introduce a family of random variables

$$\phi_{1\beta} \otimes \phi_{2\sigma} : \Omega \rightarrow \Omega_D' \times \Omega_S' \times \Omega_{2D}'$$

(1)

whose definition is

$$(\phi_{1\beta} \otimes \phi_{2\sigma})(\omega_L) = D_{1\beta}(\omega_L \times \Omega_S) \otimes D_{2\sigma}(\omega_L \times \Omega_{2D})$$

(2)

with $\omega_L \in \Omega$ and $\Omega_{1D} \in \Omega_{1D}$, and $\Omega_{2D} \in \Omega_{2D}$.

The random variable $\phi_{1\beta} \otimes \phi_{2\sigma}$ gives information on the joint outcomes that happen simultaneously at each measurement station with respect to events on the enlarged space. To make contact with Nelson’s notation, we define the notations $\sigma_{1\beta} \equiv \tilde{\sigma}_{1\beta} \times \Omega_{2D}$, $\sigma_{1\beta} = \Omega_{1D} \times \Omega_{2D}$, and $\sigma_{1\beta} = \Omega_{1D} \times \Omega_{2D}$. The events $\{\Omega_{1\beta} \in \sigma_{1\beta} \equiv \Omega_{1D} \times \Omega_{2D}\}$ contain information on the outcomes at detector 1 for axis $\beta$. We let the set of these events for all axes $\beta$ at 1 generate a sigma algebra $\mathcal{F}_1$. Then, $\mathcal{F}_1$ contains all information about the events that might happen at 1.

Similarly, we define the notations $\sigma_{2\sigma} \equiv \Omega_{1D}' \times \tilde{\sigma}_{2\sigma}$, $\sigma_{2\sigma} = \Omega_{1D}' \times \tilde{\sigma}_{2\sigma}$, and $\sigma_{2\sigma} = \Omega_{1D}' \times \Omega_{2D}'$. The events $\{\Omega_{1\sigma} \in \sigma_{1\sigma} \equiv \Omega_{1D}' \times \Omega_{2D}'\}$ contain information on the outcomes at detector 2. We let the set of these events for all axes $\sigma$ at 2 generate the sigma algebra $\mathcal{F}_2$.

The probabilities of the events that are generated by $\phi_{1\beta} \otimes \phi_{2\sigma}$ will be denoted by

$$P\left(\left\{\phi_{1\beta} \otimes \phi_{2\sigma} \in \left(\sigma_{1\beta} \cap \sigma_{2\sigma}\right)\right\}\right) \equiv P_{\phi_{1\beta} \otimes \phi_{2\sigma}}\left(\sigma_{1\beta} \cap \sigma_{2\sigma}\right).$$

(3)

The expression $P_{\phi_{1\beta} \otimes \phi_{2\sigma}}$ defines an axis dependent family of probability measures for events of the form $\sigma_{1\beta} \cap \sigma_{2\sigma}$, where

Since the family of random variables $\phi_{1\beta} \otimes \phi_{2\sigma}$ is defined on the enlarged probability space, it will give us the possibility to investigate the relationship between the outcomes at the detectors and the ones at preparation stage. We let a third sigma algebra $\mathcal{F}_3$ contain the information about the outcomes at the source. Accordingly, $\mathcal{F}_3$ is generated by the power set $\Omega_D \times P(\Omega_S) \times \Omega_{2D}$.

We can compute the conditional probability of the events at the detectors given the events in $\mathcal{F}_3$ that happen at preparation stage with:

$$P(A|\mathcal{F}_3) \equiv \text{EX}[1_A|\mathcal{F}_3]$$

(4)

Eq. (4) is the definition of a random variable for outcomes in the enlarged probability space. It can be interpreted as the revised probability of an event $A$ to happen with respect to the extra information about which events in $\mathcal{F}_3$ occur for a given outcome. Obviously, $P(A|\mathcal{F}_3)$ fulfills the following properties:

- For all events $A \in \mathcal{F}$, the random variable $P(A|\mathcal{F}_3)$ is measurable with respect to $\mathcal{F}_3$.
- We have $0 \leq P(A|\mathcal{F}_3) \leq 1$ for all events $A \in \mathcal{F}$ and outcomes on the enlarged probability space. If $A$ is the sure event $\Omega_D \times \Omega_S \times \Omega_{2D}$, the conditional probability $P(A|\mathcal{F}_3)$ is unity for all outcomes on the enlarged probability space. In case $A$ is the impossible event $\emptyset$, the random variable $P(A|\mathcal{F}_3)$ is zero for all outcomes.
- For every sequence of pairwise disjoint events $(A_n)_{n \in \mathbb{N}} \in \mathcal{F}$, the equality $P\left(\bigcup_{n=1}^{\infty} A_n|\mathcal{F}_3\right) = \sum_{n=1}^{\infty} P(A_n|\mathcal{F}_3)$ holds.
- The unconditional probability of any event $A \in \mathcal{F}$ can be computed with the expectation value $\text{EX}[P(A|\mathcal{F}_3)] = \int P(A|\mathcal{F}_3) \ dP = P(A)$.

3
For any event $A \in \mathcal{F}$ and another event $A_1 \in \mathcal{F}_S$, where $A_1 \neq \emptyset$, it follows from Eq. (3) that $P\left(\bigcap A_1 \mid \mathcal{F}_S\right) = P\left(A \mid \mathcal{F}_S\right)$ for outcomes in $A_1 \in \mathcal{F}_S$, and $P\left(\bigcap A_1 \mid \mathcal{F}_S\right) = 0$ for outcomes not in $A_1 \in \mathcal{F}_S$.

In case that the events on the enlarged probability space are generated by $\phi_1 \otimes \phi_2$, we will use the following notation:

$$P_{\phi_1 \otimes \phi_2} \left(\tilde{\sigma}_1 \times \tilde{\sigma}_2 \mid \mathcal{F}_S\right) \equiv P\left(\{\phi_1 \otimes \phi_2 \in \{\tilde{\sigma}_1 \times \tilde{\sigma}_2\}\} \mid \mathcal{F}_S\right).$$

(5)

In the analysis of the EPR experiment, we will often have to deal with equivalent events. Two events $A$ and $B$ in $\mathcal{F}$ are equivalent if

$$P \left(\bigcap A \bigcap B\right) = P(A) = P(B).$$

(6)

This means that if $A$ happens, then $B$ also happens and vice versa. This does not mean, however, that $A$ causes $B$ to happen or vice-versa. if $A$ and $B$ are equivalent to each other, then their complements $(A)^c$ and $(B)^c$ must also be equivalent to each other. Furthermore, one can show that if the event $A$ is equivalent the event $B$, and $B$ is equivalent to another event $C$, then $A$ is also equivalent to $C$. Furthermore,

Now, all the mathematical structures needed to analyse the EPR experiment have been defined and we can proceed with the necessary locality conditions.

### 2.2 Active locality

At first, Nelson defined two different forms of locality: Active locality and passive locality. The meaning of active locality is:

Whatever axes the experimenter selects at one measurement device, e.g. at 2, it does not change the outcomes at 1, as long as 1 does not lie in the future cone of 2. Active locality is a requirement from the theory of relativity. We can mathematically define active locality as follows:

The Stern Gerlach magnets may influence the outcomes of the experiment as a function of their axes. The corresponding events are generated by the axis dependent random variables $\phi_1 \otimes \phi_2$. Now we let the area of measurement station 1 be disjoint from the future cone of station 2. With the axis $\tilde{\mu}$ of 1 being left constant, different axes $\tilde{\nu} \neq \tilde{\nu}'$ of Stern-Gerlach magnet 2 are chosen. We call a theory actively local if the events $\{\phi_1 \otimes \phi_2 \in \sigma_1\} \in \mathcal{F}_1$ and $\{\phi_1 \otimes \phi_2 \in \sigma_1\} \in \mathcal{F}_1$ are equivalent. That is

$$P \left(\{\phi_1 \otimes \phi_2 \in \sigma_1\} \bigcap \{\phi_1 \otimes \phi_2 \in \sigma_1\}\right) = P_{\phi_1 \otimes \phi_2} \left(\sigma_1\right) = P_{\phi_1 \otimes \phi_2} \left(\sigma_1\right).$$

(7)

Eq. (7) implies that a change of the setting at station 2 does not modify the outcome at station 1. Note that from Eq. (7), only

$$P_{\phi_1 \otimes \phi_2} \left(\sigma_1\right) = P_{\phi_1 \otimes \phi_2} \left(\sigma_1\right)$$

(8)

is actually required by the quantum mechanical probabilities of quantum mechanics.

A similar condition should hold for the events $\{\phi_1 \otimes \phi_2 \in \sigma_2\} \in \mathcal{F}_2$ and $\{\phi_1 \otimes \phi_2 \in \sigma_2\} \in \mathcal{F}_2$ at Stern-Gerlach magnet 2, provided that the spatial region of 2 is disjoint from the future cone of magnet 1:

$$P_{\phi_1 \otimes \phi_2} \left(\sigma_2\right) = P \left(\{\phi_1 \otimes \phi_2 \in \sigma_2\} \bigcap \{\phi_1 \otimes \phi_2 \in \sigma_2\}\right) = P_{\phi_1 \otimes \phi_2} \left(\sigma_2\right).$$

(9)

In Eq. (9), we have $\tilde{\nu}' \neq \tilde{\mu}$ for the axes of detector 1, with the axis $\tilde{\nu}'$ of magnet 2 being left constant.

By Eq. (9), we can condition the probabilities of $\{\phi_1 \otimes \phi_2 \in \sigma_1\} \in \mathcal{F}_1$ and $\{\phi_1 \otimes \phi_2 \in \sigma_2\} \in \mathcal{F}_2$ with respect to $\mathcal{F}_S$ and we get

$$P_{\phi_1 \otimes \phi_2} \left(\sigma_1\mathcal{F}_S\right) = P_{\phi_1 \otimes \phi_2} \left(\sigma_1\mathcal{F}_S\right),$$

(10)

$$P_{\phi_1 \otimes \phi_2} \left(\sigma_2\mathcal{F}_S\right) = P_{\phi_1 \otimes \phi_2} \left(\sigma_2\mathcal{F}_S\right).$$

(11)
It is important to note that Eq. 8 describes the probabilities which are given by quantum mechanics. This equation forbids in any theory that describes the quantum mechanics probabilities to send instantaneous signals to spacelike separated locations. In contrast Eqs. 7 and 9 are not given by quantum mechanics and are violated by many hidden variable theories, such as Bohmian mechanics.

Because quantum mechanics does not contain Eqs. 7 and 9, it does not forbid an explanation of its outcomes with a model that violates causality. Someone who strictly insists on the relativity principle could therefore view quantum mechanics as an incomplete theory.

Furthermore, one should note that, when describing an event by \( P_{\phi_1 \oplus \phi_2} (\sigma_1 \mid \sigma_2) \), one can not conclude from active locality alone that we would have

\[
P_{\phi_1 \oplus \phi_2} (\sigma_1 \mid \sigma_2) = P_{\phi_1 \oplus \phi_2} (\sigma_2 \mid \sigma_1).
\]

The reason for this is that the event \( \{ \phi_1 \oplus \phi_2 \in \sigma_1 \} \in F_1 \) must not be equivalent to the event \( \{ \phi_1 \oplus \phi_2 \in \sigma_2 \} \in F_1 \). Similarly \( \{ \phi_1 \oplus \phi_2 \in \sigma_2 \} \in F_2 \) must not be equivalent to \( \{ \phi_1 \oplus \phi_2 \in \sigma_1 \} \in F_2 \). The physical reason for this is that spin operators, like all angular momentum operators in quantum mechanics, do not commute.

### 2.3 Passive locality

In this subsection, we consider the conditional joint probability of \( \{ \phi_1 \oplus \phi_2 \in \sigma_1 \} \in F_1 \) and \( \{ \phi_1 \oplus \phi_2 \in \sigma_2 \} \in F_2 \) with respect to \( F_S \). It gives information about events which happen simultaneously at the spatially separated locations 1 and 2. We say that passive locality holds if

\[
P_{\phi_1 \oplus \phi_2} (\sigma_1 \mid \sigma_2, F_S) = P_{\phi_1 \oplus \phi_2} (\sigma_1 \mid F_S) P_{\phi_1 \oplus \phi_2} (\sigma_2 \mid F_S),
\]

for every pair of axes \( \vec{\mu} \) and \( \vec{\nu} \). The outcomes of an experiment may be statistically dependent. Such a dependence may arise because of a preparation stage. The condition of passive locality says that the outcomes should be conditionally independent given the information about preparation stage \( F_S \). A violation of passive locality would mean that the dependence of the outcomes does not originate at \( F_S \). It is possible to have active locality without passive locality. Furthermore, a theory that violates passive locality does not have to incorporate any non-local interaction between the spatially separated measurement stations 1 and 2. However, the true physical implications of the passive locality condition will be postponed at the moment.

### 2.4 Nelson’s theorem

With the definitions above, quantum mechanics fulfills the following properties:

\[
P_{\phi_1 \oplus \phi_2} (\sigma_1 \mid \sigma_2) + P_{\phi_1 \oplus \phi_2} (\sigma_1 \mid \sigma_2) = 1,
\]

\[|E(\vec{\mu}, \vec{\nu}) - E(\vec{\mu}', \vec{\nu}) + E(\vec{\mu}, \vec{\nu}') - E(\vec{\mu}', \vec{\nu}')| > 2.
\]

Equation (14) implies that if the axes of the Stern-Gerlach magnets are the same, the spin values measured at 1 and 2 are always opposite. Using \( P_{\phi_1 \oplus \phi_2} (\sigma_1 \mid \sigma_2) = 1 \) and \( P_{\phi_1 \oplus \phi_2} (\sigma_2 \mid \sigma_1) = 1 \), we can rewrite Eq. (14) in the following form:

\[
\begin{cases}
P_{\phi_1 \oplus \phi_2} (\sigma_1 \oplus \sigma_2) = P_{\phi_1 \oplus \phi_2} (\sigma_1 \mid \sigma_2) + P_{\phi_1 \oplus \phi_2} (\sigma_2 \mid \sigma_1) = 1, \\
E(\vec{\mu}, \vec{\nu}) + E(\vec{\mu}', \vec{\nu}) + E(\vec{\mu}, \vec{\nu}') + E(\vec{\mu}', \vec{\nu}') > 2.
\end{cases}
\]

This makes clear that the events at the detectors are equivalent if the same axes at the separated measurement devices were chosen.

The function \( E(\vec{\mu}, \vec{\nu}) \) in Eq. (15) is called correlation coefficient. It is defined through:

\[
E(\vec{\mu}, \vec{\nu}) = P_{\phi_1 \oplus \phi_2} (\sigma_1 \oplus \sigma_2) + P_{\phi_1 \oplus \phi_2} (\sigma_1 \mid \sigma_2) + P_{\phi_1 \oplus \phi_2} (\sigma_2 \mid \sigma_1).
\]
It was shown by Nelson that Eq. (15) is in conflict with theories where Eq. (14) and active as well as passive locality hold. Nelson’s proof [9, 10] (with corrections in [11]) proceeds as follows:

We are interested in the properties of the conditional probabilities of the events \( \{ \phi_{1}\otimes\phi_{2}\in\sigma_{1}\} \in F_{1} \) and \( \{ \phi_{1}\otimes\phi_{2}\in\sigma_{2}\} \in F_{2} \) given the sigma algebra \( F_{S} \). At first, we will investigate, what the equivalence property of Eq. (16) combined with the assumption of passive locality implies for these conditional probabilities.

If the axes at the two detectors are the same, i.e. \( \vec{\mu} = \vec{\nu} \), we have, due to passive locality:

\[
P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\cap\sigma_{2} \right| F_{S} \right) = P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\cap\left| F_{S} \right. \right) P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{2}\left| F_{S} \right. \right).
\]

Since \( 0 \leq P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{2}\left| F_{S} \right. \right) \leq 1 \), we get

\[
P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\cap\sigma_{2} \right| F_{S} \right) \leq P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\left| F_{S} \right. \right).
\]

With the properties of the conditional probabilities stated in section 2.1, it follows that

\[
P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\left| \right. \right) = EX \left[ P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\left| F_{S} \right. \right) \right]
\]

and, similarly,

\[
P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\cap\sigma_{2} \right| F_{S} \right) = EX \left[ P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\cap\sigma_{2}\left| F_{S} \right. \right) \right].
\]

Using Eq. (16), we observe that an event \( \{ \phi_{1}\otimes\phi_{2}\in\sigma_{1}\} \) at detector 1 implies an equivalent event \( \{ \phi_{1}\otimes\phi_{2}\in\sigma_{2}\} \) at 2, i.e.

\[
P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\left| \right. \right) = P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\left| \right. \right) = P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{2}\left| \right. \right).
\]

Hence, the expectation values in Eqs. (20) and (21) are equal. Accordingly, we get with Eq. (19):

\[
P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\cap\sigma_{2} \right| F_{S} \right) = P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\left| \right. \right).
\]

and

\[
P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\cap\sigma_{2} \right| F_{S} \right) = P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{2}\left| \right. \right).
\]

Now, we will look at the consequences of active and passive locality for the conditional probabilities. Passive locality demands for events \( \{ \phi_{1}\otimes\phi_{2}\in\sigma_{1}\left| \right. \} \) at 1 and \( \{ \phi_{1}\otimes\phi_{2}\in\sigma_{2}\left| \right. \} \) at 2 that:

\[
P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\left| \right. \right) = P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\left| \right. \right) = P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{2}\left| \right. \right).
\]

Active locality implies that the actions happening at 2 can only affect anything in the future cone of 2. Since 1 does not lie in that cone, nothing what happens in 2 can affect the outcomes measured at 1. Therefore, the events \( \{ \phi_{1}\otimes\phi_{2}\in\sigma_{1}\left| \right. \} \) and \( \{ \phi_{1}\otimes\phi_{2}\in\sigma_{1}\left| \right. \} \) are equivalent:

\[
P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\left| \right. \right) = P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{1}\left| \right. \right) = P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{2}\left| \right. \right).
\]

An analogous expression is true for the events at 2:

\[
P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{2}\left| \right. \right) = P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{2}\left| \right. \right).
\]

Without loss of generality, we can select an axis \( \vec{\mu} = \vec{\nu} \) with \( \vec{\nu} \) from Eq. (27) in the Eqs. (16), (23) and (24). It then follows from Eq. (16) that the event \( \{ \phi_{1}\otimes\phi_{2}\in\sigma_{1}\left| \right. \} \) at detector 2 in the right hand side of Eq. (27) implies an equivalent event of \( \{ \phi_{1}\otimes\phi_{2}\in\sigma_{1}\left| \right. \} \) at 1. Using the Eqs. (23) and (24) with an axis \( \vec{\mu} = \vec{\nu} \), we can conclude that:

\[
P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{2}\left| \right. \right) = P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{2}\left| \right. \right) = P_{\phi_{1}\otimes\phi_{2}\in\sigma_{2}} \left( \sigma_{2}\left| \right. \right).
\]
Plugging Eq. (26) and Eq. (28) back to Eq. (25), we arrive at
\[ P_{\phi_1 \otimes \phi_2} \left( \sigma_{2\beta} = \uparrow \cap \sigma_{2\mu} = \downarrow \right| F_S \right) = P_\mu P_\nu. \] (29)

The events \( \{ \phi_{1\mu} \otimes \phi_{2\mu} \in \sigma_{2\mu} = \uparrow \} \) and \( \{ \phi_{1\nu} \otimes \phi_{2\nu} \in \sigma_{2\nu} = \downarrow \} \) are disjoint and their union is the sure event. According to section 2.1, the sum of the conditional probabilities with respect to a sigma algebra is equal to unity for such events. Therefore, we get with Eq. (28):
\[ P_{\phi_1 \otimes \phi_2} \left( \sigma_{2\beta} = \uparrow \right| F_S \right) = P_{\phi_1 \otimes \phi_2} \left( \sigma_{2\mu} = \downarrow \right| F_S \right) = 1 - P_\mu. \] (30)

Since active locality holds, the events \( \{ \phi_{1\mu} \otimes \phi_{2\mu} \in \sigma_{1\mu} = \downarrow \} \) and \( \{ \phi_{1\nu} \otimes \phi_{2\nu} \in \sigma_{1\nu} = \downarrow \} \) are equivalent, and we get, since that the complement of \( \{ \phi_{1\mu} \otimes \phi_{2\mu} \in \sigma_{1\mu} = \downarrow \} \) and \( \{ \phi_{1\nu} \otimes \phi_{2\nu} \in \sigma_{1\nu} = \downarrow \} \)

is the sure event:
\[ P_{\phi_1 \otimes \phi_2} \left( \sigma_{1\mu} = \downarrow \right| F_S \right) = P_{\phi_1 \otimes \phi_2} \left( \sigma_{1\nu} = \downarrow \right| F_S \right) = 1 - P_\nu. \] (31)

Due to passive locality and the Eqs. (31) and (30), we have for the events \( \{ \phi_{1\mu} \otimes \phi_{2\mu} \in \sigma_{1\mu} = \downarrow \} \) and \( \{ \phi_{1\nu} \otimes \phi_{2\nu} \in \sigma_{1\nu} = \downarrow \} \):
\[ P_{\phi_1 \otimes \phi_2} \left( \sigma_{1\mu} = \downarrow \cap \sigma_{1\nu} = \downarrow \right| F_S \right) = (1 - P_\mu)(1 - P_\nu). \] (32)

In the same way, we can derive the relations
\[ P_{\phi_1 \otimes \phi_2} \left( \sigma_{1\mu} = \uparrow \cap \sigma_{1\nu} = \uparrow \right| F_S \right) = P_\mu (1 - P_\nu) \] (33)

and
\[ P_{\phi_1 \otimes \phi_2} \left( \sigma_{1\mu} = \uparrow \cap \sigma_{1\nu} = \downarrow \right| F_S \right) = (1 - P_\mu) P_\nu. \] (34)

Using the Eqs. (33), (34), (29) and (32), we may define the function
\[ E \left( \bar{\mu}, \bar{\nu} \right| F_S \right) \equiv P_\mu (1 - P_\nu) + (1 - P_\mu) P_\nu - P_\mu P_\nu - (1 - P_\mu)(1 - P_\nu). \] (35)

The conditional probabilities in Eq. (35) are all in the range \( 0 \leq P_{\phi_1 \otimes \phi_2} \left( \sigma_{1\mu} \cap \sigma_{2\nu} \right| F_S \right) \leq 1 \). Therefore, the following inequality can be computed with four arbitrary axes \( \bar{\mu}, \bar{\mu}', \bar{\nu}, \bar{\nu}' \) at the two stations 1 and 2:
\[ |E \left( \bar{\mu}, \bar{\nu} \right| F_S \right) + E \left( \bar{\mu}', \bar{\nu}' \right| F_S \right) + E \left( \bar{\mu}', \bar{\nu} \right| F_S \right) - E \left( \bar{\mu}, \bar{\nu}' \right| F_S \right) \leq 2. \] (36)

If we replace the conditional probabilities in Eq. (36) by their corresponding unconditional probability measures, we arrive at the correlation coefficient from Eq. (15). The unconditional probabilities are also in the range of \( 0 \leq P_{\phi_1 \otimes \phi_2} \left( \sigma_{1\mu} \cap \sigma_{2\nu} \right) \leq 1 \), and they are given by the expectation values of the conditional probabilities. Hence, an inequality analogous to Eq. (35) must be true for them:
\[ |E \left( \bar{\mu}, \bar{\nu} \right) + E \left( \bar{\mu}', \bar{\nu}' \right) + E \left( \bar{\mu}', \bar{\nu} \right) - E \left( \bar{\mu}, \bar{\nu}' \right) | \leq 2. \] (37)

Equation (37) is called Clauser-Holt-Shimony-Horne inequality [8]. It is a version of the statement in Bell’s first article [9]
\[ |E \left( \bar{\mu}, \bar{\nu} \right) - E \left( \bar{\mu}, \bar{\nu}' \right) | \leq 1, \] (38)

which can be similarly derived. Both Eq. (37) and Eq. (38) are violated in quantum mechanics and this violation was confirmed experimentally in 1982 [20]. It should be noted that Jarret [21] arrived at a similar conclusion even though he did not formulate his article within rigorous probability theory.
3 Implications of passive locality

We begin by analysing the consequences of passive locality. Faris showed in [13] that passive locality, when combined with relation (14) from quantum mechanics, immediately leads to another condition, which he calls “deterministic passive locality”. According to Faris, Deterministic passive locality, when combined with relation (14) from quantum mechanics, immediately leads to another

Let the events \( \{ \phi_1 \otimes \phi_2 \in \sigma_1 \} \) at 1 and \( \{ \phi_1 \otimes \phi_2 \in \sigma_2 \} \) at 2 be equivalent with respect to \( P_{\phi_1 \otimes \phi_2} \) and passive locality hold. Then, there must be an event \( A_{1S} \) at preparation stage, which is equivalent to both \( \{ \phi_1 \otimes \phi_2 \in \sigma_1 \} \) at 1 and \( \{ \phi_1 \otimes \phi_2 \in \sigma_2 \} \) at 2.

In his contribution, Faris states that a similar result is presented by Redhead in [23] at pp. 101-102. Redhead claims, it would have been discovered at first by Suppes and Zanotti [24].

The derivation below will follow closely the lines of Faris:

As shown in section 2.4, it results from the equivalence property of Eq. (16) and the passive locality condition of Eq. (13) that

\[
P_{\phi_1 \otimes \phi_2} \left( \sigma_1 \bigcap \sigma_2 \right| F_S) = P_{\phi_1 \otimes \phi_2} (\sigma_1|F_S) \tag{39}
\]

and, similarly,

\[
P_{\phi_1 \otimes \phi_2} \left( \sigma_1 \bigcap \sigma_2 \right| F_S) = P_{\phi_1 \otimes \phi_2} (\sigma_2|F_S). \tag{40}
\]

Using the assumption of passive locality again, we get with Eq. (39) and Eq. (40):

\[
P_{\phi_1 \otimes \phi_2} (\sigma_1|F_S) = P_{\phi_1 \otimes \phi_2} (\sigma_2|F_S) = P_{\phi_1 \otimes \phi_2} (\sigma_1|F_S) P_{\phi_1 \otimes \phi_2} (\sigma_2|F_S) = (P_{\phi_1 \otimes \phi_2} (\sigma_1|F_S))^2. \tag{41}
\]

Eq. (41) implies, that the random variable \( P_{\phi_1 \otimes \phi_2} (\sigma_1|F_S) \) can only have the values 1 and 0. The union event of the outcomes in the enlarged probability space for which \( P_{\phi_1 \otimes \phi_2} (\sigma_1|F_S) \) has the value 1 will be denoted by \( A_{1S} \). From this definition and Eq. (41), it follows that \( P_{\phi_1 \otimes \phi_2} (\sigma_1|F_S) \) is equal to the indicator function of \( A_{1S} \):

\[
P_{\phi_1 \otimes \phi_2} (\sigma_1|F_S) = 1_{A_{1S}}. \tag{42}
\]

According to section 2.1, \( P_{\phi_1 \otimes \phi_2} (\sigma_1|F_S) \) is measurable with respect to \( F_S \). Therefore, we must have \( A_{1S} \in F_S \).

Now, we recall the definition

\[
P_{\phi_1 \otimes \phi_2} (\sigma_1|F_S) = \frac{1}{\text{EX}[P_{\phi_1 \otimes \phi_2} (\sigma_1|F_S)]} \tag{43}
\]

from section 2.1. In Eq. (20), the unconditional probability of \( \{ \phi_1 \otimes \phi_2 \in \sigma_1 \} \) is given by the expectation value

\[
P_{\phi_1 \otimes \phi_2} (\sigma_1|F_S) = \text{EX} \left[ P_{\phi_1 \otimes \phi_2} (\sigma_1|F_S) \right] \tag{44}
\]

Similarly, the probability of an event \( A_{1S} \) on the enlarged probability space with its measure \( P \) can be computed by the expectation value of the indicator function \( 1_{A_{1S}} \):

\[
\text{EX}[1_{A_{1S}}] = \int 1_{A_{1S}} dP = P(A_{1S}) \tag{45}
\]

Using the Eqs. (20), (19), (22) and (45), we can conclude that

\[
P_{\phi_1 \otimes \phi_2} (\sigma_1|F_S) = P(A_{1S}) \tag{46}
\]
In section 2.1, we also have learnt that in case of $A_{1S} \neq \emptyset$, the conditional probability of the intersection between the two events $\{\phi_{1\mu} \otimes \phi_{2\mu} \in \sigma_{1\mu}\}$ and $A_{1S} \in \mathcal{F}_S$ is equal to

\[
P \left( \{\phi_{1\mu} \otimes \phi_{2\mu} \in \sigma_{1\mu}\} \cap A_{1S} \right| \mathcal{F}_S) = \begin{cases} P \left( \{\phi_{1\mu} \otimes \phi_{2\mu} \in \sigma_{1\mu}\} | \mathcal{F}_S \right) & \text{for outcomes in } A_{1S}, \\ 0 & \text{for outcomes not in } A_{1S}. \end{cases}
\]

(47)

On the other hand, we get with $A_{1S} = \emptyset$:

\[
P \left( \{\phi_{1\mu} \otimes \phi_{2\mu} \in \sigma_{1\mu}\} \cap A_{1S} \right| \mathcal{F}_S) = P(\emptyset | \mathcal{F}_S) = 0 \text{ for all outcomes.}
\]

(48)

The Eqs. (47), (43) and (42) lead to the relation

\[
P \left( \{\phi_{1\mu} \otimes \phi_{2\mu} \in \sigma_{1\mu}\} \cap A_{1S} \right| \mathcal{F}_S) = P(\emptyset | \mathcal{F}_S) = 0 \text{ for all outcomes.}
\]

(49)

Due to Eq. (48) and the definition of $A_{1S}$, Eq. (49) also holds in case of $A_{1S} = \emptyset$. Computing the expectation value from both sides of Eq. (41) yields

\[
P \left( \{\phi_{1\mu} \otimes \phi_{2\mu} \in \sigma_{1\mu}\} \cap A_{1S} \right) = P(A_{1S}).
\]

(50)

Eqs. (40) and (50) imply that the event $\{\phi_{1\mu} \otimes \phi_{2\mu} \in \sigma_{1\mu}\} \in \mathcal{F}_1$ is equivalent to an event $A_{1S} \in \mathcal{F}_S$. It follows from Eq. (50), Eq. (40) and the second line of Eq. (41) that

\[
P_{\phi_{1\mu} \otimes \phi_{2\mu}} (\sigma_{1\mu} | \mathcal{F}_S) = (P_{\phi_{1\mu} \otimes \phi_{2\mu}} (\sigma_{2\mu} | \mathcal{F}_S))^2.
\]

(51)

Using Eq. (42) together with Eq. (51), we arrive at

\[
P_{\phi_{1\mu} \otimes \phi_{2\mu}} (\sigma_{1\mu} | \mathcal{F}_S) = P_{\phi_{1\mu} \otimes \phi_{2\mu}} (\sigma_{2\mu} | \mathcal{F}_S) = 1_{A_{1S}}.
\]

(52)

If an event $A \in \mathcal{F}$ is equivalent to an event $B \in \mathcal{F}$ and $B$ is equivalent to another event $C \in \mathcal{F}$, then $A$ is also equivalent to $C$. The events $\{\phi_{1\mu} \otimes \phi_{2\mu} \in \sigma_{1\mu}\} \in \mathcal{F}_1$, $\{\phi_{1\mu} \otimes \phi_{2\mu} \in \sigma_{2\mu}\} \in \mathcal{F}_2$ and $A_{1S} \in \mathcal{F}_S$ are not only in their sub-sigma algebras but also in $\mathcal{F}$. For this reason, the event $\{\phi_{1\mu} \otimes \phi_{2\mu} \in \sigma_{2\mu}\} \in \mathcal{F}_2$ is equivalent to the same event $A_{1S} \in \mathcal{F}_S$ which $\{\phi_{1\mu} \otimes \phi_{2\mu} \in \sigma_{1\mu}\} \in \mathcal{F}_1$ is equivalent to.

So we have the result that passive locality together with the exact anti correlations of entangled spin states imply that deterministic passive locality holds. The latter states that the events at the detectors must be predetermined by another event at preparation stage.

4 The relation between Bell’s theorem and the Free Will theorem of Conway and Kochen

In this section, we review a Proof of Bell’s inequalities that is given by Faris in [13]. It is interesting, because it relates Bell’s theorem to what Kochen and Conway call the Free Will theorem [27].

In the previous section, we have shown that the event $\{\phi_{1\mu} \otimes \phi_{2\mu} \in \sigma_{1\mu} = \uparrow\} \in \mathcal{F}_1$ for a pair of axes $\vec{A}$ and $\vec{B}$ is equivalent to some event $A_{1S} \in \mathcal{F}_S$ that is generated by the random variable $\chi_{\vec{A}} := P_{\phi_{1\mu} \otimes \phi_{2\mu}} (\sigma_{1\mu} | \mathcal{F}_S)$.

The active locality conditions of (7) and (9) then imply that

\[
P_{\phi_{1\mu} \otimes \phi_{2\mu}} (\sigma_{1\mu} = \uparrow) = P_{\phi_{1\mu} \otimes \phi_{2\mu}} (\sigma_{1\mu} = \uparrow) = P \left( \{\phi_{1\mu} \otimes \phi_{2\mu} \in (\sigma_{1\mu} = \uparrow)\} \cap \{\chi_{\vec{A}} \in A_{1S}\} \right)
\]

\[
= P \left( \{\chi_{\vec{A}} \in A_{1S}\} \right)
\]

\[
= P \left( \{\chi_{\vec{A}} \in A_{1S}\} \right)
\]

\[
= P \chi_{\vec{A}} (A_{1S}).
\]

(53)
The events $\{\chi_{A\bar{B}} \in A_1 S\}$ happen at preparation stage. Therefore, we have to restrict the possible axis dependence of that event even more. If we take the active locality condition seriously, then, the event $\{\chi_{A\bar{B}} \in A_1 S\}$ should indeed be equivalent to an event $\{\chi_{\bar{A}B'} \in A_1 S\}$, for arbitrary axes $\bar{A} \neq \bar{A}$ and $B' \neq \bar{B}$.

Here, the conflict with quantum mechanics arises. Since angular momentum operators do not commute, there can, at least in quantum mechanics, not be a spin outcome for axis $\bar{A}$ that is equivalent to another spin outcome for an arbitrary axis $\bar{A}' \neq \bar{A}$ at the same detector.

In Faris proof from section 3, passive locality together with the exact anticorrelations result in determinism. Now determinism together with active locality says that even for a set of non-compatible observables, there must be an event that determines the values of these observables for all times. We will now proceed with Faris Proof of Bell's theorem and show, that this is not possible.

A similar argument as given above implies that the event $\{\phi_{1\bar{B}} \otimes \phi_{2\bar{B}} \in (\sigma_{2\bar{B}} = \downarrow)\} \in F_2$ for an axis $\bar{B}$ is equivalent to an event $B_{2S} \in F_S$ and we get with active locality:

$$
P_{\phi_{1\bar{B}} \otimes \phi_{2\bar{B}}}(\sigma_{2\bar{B}} = \downarrow) = P\left(\left\{\phi_{1\bar{B}} \otimes \phi_{2\bar{B}} \in (\sigma_{2\bar{B}} = \downarrow)\right\} \cap \{\chi_{\bar{B}\bar{B}} \in B_{1S}\}\right)
= P\left(\left\{\phi_{1\bar{A}} \otimes \phi_{2\bar{A}} \in (\sigma_{2\bar{A}} = \downarrow)\right\} \cap \{\chi_{\bar{A}\bar{A}} \in B_{1S}\}\right)
= P_{\chi_{\bar{A}\bar{A}}}(B_{2S})
= P_{\chi_{\bar{A}\bar{B}}}(B_{2S}).
$$

(54)

Since $\{\phi_{1\bar{B}} \otimes \phi_{2\bar{B}} \in (\sigma_{1\bar{B}} = \downarrow)\}$ is the complement of $\{\phi_{1\bar{B}} \otimes \phi_{2\bar{B}} \in (\sigma_{1\bar{B}} = \uparrow)\}$, we also have

$$
P_{\phi_{1\bar{B}} \otimes \phi_{2\bar{B}}}(\sigma_{1\bar{B}} = \uparrow \cap \sigma_{2\bar{B}} = \uparrow) = P\left(\{\chi_{\bar{A}\bar{B}} \in A_1 S\} \cap \{\chi_{\bar{A}\bar{B}} \in (B_{2S})^C\}\right)
= P_{\chi_{\bar{A}\bar{B}}}(A_1 S \cap (B_{2S})^C)
= P_{\chi_{\bar{A}\bar{B}'}}(A_1 S \cap (B_{2S})^C),
$$

(55)

where $(B_{2S})^C$ denotes the complement of $B_{2S}$.

We can do this for three axes at the detectors $\bar{A}, \bar{B}, \bar{C}$.

$$
P_{\phi_{1\bar{A}} \otimes \phi_{2\bar{B}}}(\sigma_{1\bar{A}} = \uparrow \cap \sigma_{2\bar{B}} = \uparrow) + P_{\phi_{1\bar{B}} \otimes \phi_{2\bar{C}}}(\sigma_{1\bar{B}} = \uparrow \cap \sigma_{2\bar{C}} = \uparrow)
+ P_{\phi_{1\bar{C}} \otimes \phi_{2\bar{A}}}(\sigma_{1\bar{C}} = \uparrow \cap \sigma_{2\bar{A}} = \uparrow)
= P_{\chi_{\bar{A}\bar{B}}}(A_1 S \cap (B_{2S})^C) + P_{\chi_{\bar{B}\bar{C}}}(B_{1S} \cap (C_{2S})^C)
+ P_{\chi_{\bar{C}\bar{A}}}(C_{1S} \cap (A_{2S})^C)
$$

(56)

The last equation contains the axes $\bar{A}', \bar{B}', \bar{C}'$. These axes are arbitrary and by active locality the computed events at preparation stage are equivalent for all pairs of axes. Therefore, we can set $\bar{A}' = B' = C'$.

$$
P_{\phi_{1\bar{A}} \otimes \phi_{2\bar{B}}}(\sigma_{1\bar{A}} = \uparrow \cap \sigma_{2\bar{B}} = \uparrow) + P_{\phi_{1\bar{B}} \otimes \phi_{2\bar{C}}}(\sigma_{1\bar{B}} = \uparrow \cap \sigma_{2\bar{C}} = \uparrow)
+ P_{\phi_{1\bar{C}} \otimes \phi_{2\bar{A}}}(\sigma_{1\bar{C}} = \uparrow \cap \sigma_{2\bar{A}} = \uparrow)
= P_{\chi_{\bar{A}\bar{B}}}(A_1 S \cap (B_{2S})^C) + P_{\chi_{\bar{B}\bar{C}}}(B_{1S} \cap (C_{2S})^C)
+ P_{\chi_{\bar{C}\bar{A}}}(C_{1S} \cap (A_{2S})^C)
$$

(57)

For the event $(A_{2S})^C$, there is an equivalent event $\{\phi_{1\bar{A}} \otimes \phi_{2\bar{A}} \in (\sigma_{2\bar{A}} = \uparrow)\}$ at the detectors. Now this event is, by Eq. (15), equivalent to $\{\phi_{1\bar{A}} \otimes \phi_{2\bar{A}} \in (\sigma_{1\bar{A}} = \downarrow)\}$, which, by active and passive locality, must be equivalent to the event $(A_1 S)^C \in F_S$. Making the substitutions with the events
(A_{2S})^C, (B_{2S})^C, and (C_{2S})^C, we arrive at:

\begin{align*}
P_{\phi_1\hat{A} \otimes \phi_2\hat{B}} (\sigma_1 = \uparrow \cap \sigma_2 \neq \uparrow) + P_{\phi_1\hat{B} \otimes \phi_2\hat{C}} (\sigma_1 = \uparrow \cap \sigma_2 \sigma_3 = \uparrow) \\
+ P_{\phi_1\hat{C} \otimes \phi_2\hat{A}} (\sigma_1 \neq \uparrow \cap \sigma_2 \sigma_3 = \uparrow) \\
= P_{\chi_{\hat{A} \otimes \hat{B}}} (A_{1S} \cap (B_{1S})^C) + P_{\chi_{\hat{A} \otimes \hat{C}}} (B_{1S} \cap (C_{1S})^C) \\
+ P_{\chi_{\hat{B} \otimes \hat{C}}} (C_{1S} \cap (A_{1S})^C) \\
\leq 1 \tag{58}
\end{align*}

In the last step, we have used that the events whose probabilities are computed are exclusive.

Now, one can imagine a theory, where the events at the detectors do not depend on the measurement devices at all. In such a theory, we would have

\begin{align*}
P_{\phi_1\hat{A} \otimes \phi_2\hat{B}} (\sigma_1 = \uparrow \cap \sigma_2 \neq \uparrow) + P_{\phi_1\hat{B} \otimes \phi_2\hat{C}} (\sigma_1 = \uparrow \cap \sigma_2 \sigma_3 = \uparrow) \\
+ P_{\phi_1\hat{C} \otimes \phi_2\hat{A}} (\sigma_1 \neq \uparrow \cap \sigma_2 \sigma_3 = \uparrow) \\
= P_{\phi_1\hat{A} \otimes \phi_2\hat{B}} (\sigma_1 = \uparrow \cap \sigma_2 \neq \uparrow) + P_{\phi_1\hat{B} \otimes \phi_2\hat{A}} (\sigma_1 = \uparrow \cap \sigma_2 \sigma_3 = \uparrow) \\
+ P_{\phi_1\hat{C} \otimes \phi_2\hat{A}} (\sigma_1 \neq \uparrow \cap \sigma_2 \sigma_3 = \uparrow) \\
\leq 1 \tag{59}
\end{align*}

Since no outcome is assumed to depend on an axis, we can substitute the events at 2 by their equivalent events at 1:

\begin{align*}
P_{\phi_1\hat{A} \otimes \phi_2\hat{B}} (\sigma_1 = \uparrow \cap \sigma_2 \neq \uparrow) + P_{\phi_1\hat{B} \otimes \phi_2\hat{C}} (\sigma_1 = \uparrow \cap \sigma_2 \sigma_3 = \uparrow) \\
+ P_{\phi_1\hat{C} \otimes \phi_2\hat{A}} (\sigma_1 \neq \uparrow \cap \sigma_2 \sigma_3 = \uparrow) \\
= P_{\phi_1\hat{A} \otimes \phi_1\hat{A}} (\sigma_1 = \uparrow \cap \sigma_1 \neq \uparrow) + P_{\phi_1\hat{B} \otimes \phi_1\hat{A}} (\sigma_1 = \uparrow \cap \sigma_1 \sigma_3 = \uparrow) \\
+ P_{\phi_1\hat{C} \otimes \phi_1\hat{A}} (\sigma_1 \neq \uparrow \cap \sigma_1 \sigma_3 = \uparrow) \\
\leq 1 \tag{60}
\end{align*}

And we arrive again at Bell’s inequality.

Hence, there can be no theory for the singlet state at all, where the outcomes do not depend on the settings of the detectors.

There may be theories for the singlet state, where an event at preparation stage determines the outcomes at the detectors. This preparation event then would have to depend on the settings of the measurement devices but these devices can later be chosen by the experimenters at will.

Recently, Conway and Kochen discovered this result independently and called it Free-Will theorem. If active locality holds and the experimenters can choose their axis freely, the outcomes cannot be pre-determined by a preparation event. As a result, the particles must behave randomly.

(From Nelson’s writings on Bell’s theorem, it seems that Nelson knew this since 1984. Apparently, Nelson’s writings were unknown even to Kochen, who is in the same mathematics department at Princeton.)

The author of this note believes that experimenters can choose their axis freely and active locality holds. Therefore, and because non-local theories can never be written manifestly covariant, the author thinks that non-local models of quantum mechanics should not be pursued.

On his article, Faris summarizes his proof of Bell’s theorem as follows:

The logic of the proof is elementary but subtle, [...] The first part of the argument is the observation that when the two magnetic field gradients are taken in opposite direction, the results exactly coincide. The deterministic passive locality assumption says that this implies that the randomness must have been introduced at the time the particles were prepared, since a later source of randomness would spoil the coincidence. Thus, with the magnetic field gradient in opposite direction, each spin result is determined by...
something that happened at the preparation stage. The spins may also be measured in the situation when the magnetic field gradient is not in the opposite direction to the magnetic field gradient for the other particle. According to active locality, the results of spin measurements should not depend on the direction used for the measurement of the other particle. Therefore, for the magnetic field gradients in arbitrary directions, the spin results for each particle are still determined by something that happened at the preparation stage. So, the magnets are not responsible for the randomness, it must be intrinsic to the system of two particles. In case of magnetic field gradients that are not the same or opposite, the dependence is strong but not perfect. This dependence is entirely due to something that happened at the preparation stage. The magnet configurations do not affect the probabilities. In this situation, [Eq. 58] says that the dependence is strong enough so that there can’t be an intrinsic overall outcome that determines the result of each measurement.

5 Conclusions

The analysis of Nelson and Faris make clear that Bell’s theorem does not require a theory that describes the quantum mechanical correlations with random variables to be non-local. In the words of a recent article of Nelson [25]:

In quantum mechanics, if there are two dynamically uncoupled systems, an alteration of the second system in no way affects the first, even if the two systems are entangled.

However, in order to maintain locality, a theory must give up determinism, since latter is a consequence of passive locality.

In the words of Faris:

A violation of active locality would be upsetting, since it would mean that an active intervention at one point could influence the outcome at distant points. However, a violation of passive locality would only mean that dependence between simultaneous events at distant locations need have no explanation in terms of prior events. This is not clearly ruled out, but it is not evident how to construct such a theory.

Nelson’s analysis of Bell’s theorem also shows that, unfortunately, quantum mechanics contains no principle which actually forbids non-local theories. This is not what one should expect from a theory that can be written covariantly, and it is the reason for the interpretational problems of quantum mechanics. To construct an (actively) local theory of random variables that produces the quantum mechanical outcomes is therefore a reasonable scientific program. Only if we have such a model, we are able to exclude the various non-local and therefore non-physical interpretations of quantum theory.

Nelson’s stochastic mechanics could, for example, be developed into a local field theory that violates passive locality for bosons [9, 10], including the electromagnetic field [26]. It was possible to quantise the linearized gravitational field from it [22]. Perhaps, a local theory of random variables that underlies quantum mechanics could be applied to field theoretical problems in strong gravitational fields, where ordinary quantum mechanics is not applicable. This is certainly worthwhile to investigate.

6 Relation between the assumptions behind Bell’s inequality and the locality principle of Einstein

Finally, it is interesting to ask whether EPR would have called theories that violate passive locality to be complete. In their article [1], they write:

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 reality corresponding to this physical quantity.
As we have shown in section 3, this reality criterion can not be applied to any theory that violates the passive locality condition of Eq. (13). Furthermore, as there can be no device independent model for the singlett state at all, even in a deterministic theory, the experimenter at 2 can not predict the outcome at 1, as he does not know the setting at 1, on which the outcome at 1 must depend. In their article, EPR do not write whether they call a theory, where one is unable to predict the outcomes with certainty, as realistic or not.

Einstein seemed to notice that the reality condition in [1] would be problematic. In a letter to Schrödinger, [28], Einstein complained that the main point which he wanted to make would have been "buried by erudition". Einstein then wrote own articles on his views about quantum mechanics. In his own writings, the reality criterion becomes a quite different one. He writes 1948 in a letter to Born:

I just want to explain what I mean when I say that we should try to hold on to physical reality. We are, to be sure, all of us aware of the situation regarding what will turn out to be the basic foundational concepts in physics: the point-mass or the particle is surely not among them; the field, in the Faraday-Maxwell sense, might be, but not with certainty. But that which we conceive as existing (real) should somehow be localized in time and space. That is, the real in one part of space, A, should (in theory) somehow exist independently of that which is thought of as real in another part of space, B. If a physical system stretches over the parts of space A and B, then what is present in B should somehow have an existence independent of what is present in A. What is actually present in B should thus not depend upon the type of measurement carried out in the part of space, A; it should also be independent of whether or not, after all, a measurement is made in A. If one adheres to this program, then one can hardly view the quantum-theoretical description as a complete representation of the physically real. If one attempts, nevertheless, so to view it, then one must assume that the physically real in B undergoes a sudden change because of a measurement in A.

Clearly, the notion of "independence" here is not in the sense of statistical independence, or the "conditional independence" that is described by the passive locality condition of Eq. (13). What Einstein meant is more in the sense of the active locality condition

\[
P(\{\phi_{1\mu} \otimes \phi_{2\nu} \in \sigma_{1\mu}\} \cap \{\phi_{1\mu} \otimes \phi_{2\nu'} \in \sigma_{1\mu}\}) = P_{\phi_{1\mu} \otimes \phi_{2\nu}}(\sigma_{1\mu}) = P_{\phi_{1\mu} \otimes \phi_{2\nu'}}(\sigma_{1\mu}),
\]

which does not have to be violated in a random variable model of the singlett state. Einstein wrote the letter quoted above in 1948. It took time until 1986, when Edward Nelson discovered that exactly this locality principle which Einstein wanted to implement into quantum mechanics, was not forbidden by any theorem. On the contrary, since quantum mechanics just requires

\[
P_{\phi_{1\mu} \otimes \phi_{2\nu}}(\sigma_{1\mu}) = P_{\phi_{1\mu} \otimes \phi_{2\nu'}}(\sigma_{1\mu})
\]

, it does not say anything on whether active locality holds or not. So the construction of a model that contains this locality principle is still an open problem.

References

[1] A. Einstein, B. Podolski, N. Rosen, Phys. Rev. 47, 777 (1935)
[2] D. Bohm, Quantum Theory (Prentice-Hall, Englewood Cliffs, New Jersey, 1951), pp. 614-619
[3] D. Bohm, Y. Aharanov, Phys. Rev. 108, 1070 (1957)
[4] J. S. Bell, The Theory of Local Beables, in: [7], p. 52
[5] J. S. Bell, On The Einstein Podolsky Rosen Paradox, in: [7], p. 14
[6] J. S. Bell, Introduction to the hidden-variable question, in: [7], p. 29
[7] J. S. Bell,Speakable and Unspeakable in Quantum Mechanics (Cambridge University Press, Cambridge, 1987)
[8] J. F. Clauser, R. A. Holt, M. A. Horne, A. Shimony, Phys. Rev. Lett. 23, 880 (1969)
[9] E. Nelson, Field Theory and the Future of Stochastic Mechanics, in: Stochastic Processes in Classical and Quantum Systems, Proc. Ascona, Switzerland, 1985, edited by S. Albeverio, G. Casati and D. Merlini, Lecture Notes in Physics 262 (Springer, Berlin, 1988), p. 438
[10] E. Nelson, The Locality Problem in Stochastic Mechanics, in: New Techniques and Ideas in Quantum Measurement Theory, edited by D. Greenberger, Ann. N. Y. Acad. Sci. 480, 533 (1986)
[11] E. Nelson, Stochastic Mechanics and Random Fields, in: École d’Été de Probabilités de Saint Flour XV-XVII 1985-87, Lecture Notes in Mathematics 1362 (Springer, Berlin, 1988), p. 427
[12] N. D. Mermin, Am. J. Phys. 49, 940 (1981)
[13] W. G. Faris, Probability in Quantum Mechanics, appendix to: D. Wick, The Infamous Boundary: Seven Decades of Controversy in Quantum Physics (Birkhauser, Boston, 1995)
[14] E. Nelson, Phys. Rev. 150, 1079 (1966)
[15] E. Nelson, Quantum Fluctuations (Princeton University Press, Princeton, New Jersey, 1985)
[16] P. Billingsley, Probability and Measure, 3rd edition (Wiley-Interscience, New York, 1995)
[17] R. M. Dudley, Real Analysis and Probability, 2nd edition (Cambridge University Press, Cambridge, 2002)
[18] H. Bauer, Maß- und Integrationstheorie, 2. Auflage (De Gruyter, Berlin, 1998)
[19] H. Bauer, Wahrscheinlichkeits-theorie, 5. Auflage (De Gruyter, Berlin, 2001)
[20] A. Aspect, J. Dalibard, G. Roger, Phys. Rev. Lett. 49, 1804 (1982)
[21] J. P. Jarret, Noûs 18, 569 (1984)
[22] M. Davidson, J. Math. Phys. 23, 132 (1982), [arXiv:quant-ph/0112076]
[23] M. Redhead, Incompleteness, Nonlocality and Realism (Clarendon Press, Oxford, 1987)
[24] P. Suppes, M. Zanotti, On the Determinism of Hidden Variable Theories with Strict Correlation and Conditional Statistical Independence of Observables, in: Logic and Probability in Quantum Mechanics, edited by P. Suppes (Reidel, Dordrecht, 1976), p. 445
[25] E. Nelson, J. Phys.: Conf. Ser. 361 012011, 2012
[26] S. C. Lim, Phys. Rev. D 32, 13841388 (1985)
[27] J. Conway, S. Kochen, Found. Phys. 36, 1441 (2006), [arXiv:quant-ph/0604079]
[28] D. Howard, Einstein on Locality and Separability, Studies in History and Philosophy of Science 16 (1985), 171-201.
[29] Born, Max, The Born Einstein Letters (Macmillan Press, 1971), p. 162 Nymphenburger.