Bell's theorem: Critique of proofs with and without inequalities

Karl Hess and Walter Philipp

Citation: AIP Conference Proceedings **750**, 150 (2005); doi: 10.1063/1.1874568
View online: http://dx.doi.org/10.1063/1.1874568
View Table of Contents: http://scitation.aip.org/content/aip/proceeding/aipcp/750?ver=pdfcov
Published by the AIP Publishing

Bell's theorem: Critique of proofs with and without inequalities

Karl Hess* and Walter Philipp†

*Beckman Institute, Department of Electrical and Computer Engineering and Department of Physics,University of Illinois, Urbana, Il 61801
†Beckman Institute, Department of Statistics and Department of Mathematics, University of Illinois,Urbana, Il 61801

Abstract. Most of the standard proofs of the Bell theorem are based on the Kolmogorov axioms of probability theory. We show that these proofs contain mathematical steps that cannot be reconciled with the Kolmogorov axioms. Specifically we demonstrate that these proofs ignore the conclusion of a theorem of Vorob'ev on the consistency of joint distributions. As a consequence Bell's theorem stated in its full generality remains unproven, in particular, for extended parameter spaces that are still objective local and that include instrument parameters that are correlated by both time and instrument settings. Although the Bell theorem correctly rules out certain small classes of hidden variables, for these extended parameter spaces the standard proofs come to a halt. The Greenberger-Horne-Zeilinger (GHZ) approach is based on similar fallacious arguments. For this case we are able to present an objective local computer experiment that simulates the experimental test of GHZ performed by Pan, Bouwmeester, Daniell, Weinfurter and Zeilinger and that directly contradicts their claim that Einstein-local elements of reality can neither explain the results of quantum mechanical theory nor their experimental results.

INTRODUCTION

Consider three joint pair probability distributions defined on the Euclidean plane such that any two of these three joint pair distributions share a common marginal. Let these three pair distributions be generated by the following pairs of random variables (A,B), (A,C) and (B,C). Then, according to a theorem of Vorob'ev [1] it may not be possible to realize these three probability distributions on a common probability space in the following sense. It may not be possible to find on any probability space three random variables A,B,C (and a corresponding distribution formed for that triple) with the property that the three joint pair distributions that can be formed from that triple will coincide with the initially given joint pair distributions for (A,B), (A,C) and (B,C). Here is a modification of the example that Vorob'ev used as the opening statement for his paper. As before we label the three joint distributions in terms of pairs of random variables (A,B), (A,C), and (B,C). All random variables assume only the values +1 and -1. We define these three joint probability distributions according to the following table:

TABLE 1. Modified example of Vorob'ev [1].

	$(+1,+1)$	$(+1,-1)$	$(-1,+1)$	$(-1,-1)$
(A, B)	$\frac{1}{4}(1+\frac{1}{\sqrt{2}})$	$\frac{1}{4}(1-\frac{1}{\sqrt{2}})$	$\frac{1}{4}(1-\frac{1}{\sqrt{2}})$	$\frac{1}{4}(1+\frac{1}{\sqrt{2}})$
(A, C)	$\frac{1}{4}(1+\frac{1}{\sqrt{2}})$	$\frac{1}{4}(1-\frac{1}{\sqrt{2}})$	$\frac{1}{4}(1-\frac{1}{\sqrt{2}})$	$\frac{1}{4}(1+\frac{1}{\sqrt{2}})$
(B, C)	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

Then it is easy to see that it is not possible to assign a non-negative probability to the event $(A=-1,B=-1,C=-1)$ consistent with Table 1. Indeed, this latter probability could not be more than $\frac{1}{4}$ because it cannot exceed $P(B=-1,C=-1)=\frac{1}{4}$; similarly $P(A=-1,B=-1,C=+1)\leq P(A=-1,C=+1)=\frac{1}{4}(1-\frac{1}{\sqrt{2}})$. Adding these two probabilities we obtain $P(A=-1,B=-1)\leq\frac{1}{4}(2-\frac{1}{\sqrt{2}})$ a bound smaller than the value assigned in Table 1. Thus A,B,C *can not be defined on a common probability space* such that the joint distribution of any of the three pairs that possibly could be formed from them coincides with the pair distributions defined by Table 1. Note that any two of the three

CP750, *Foundations of Probability and Physics - 3,* edited by A. Khrennikov

joint distributions share the same marginal, namely each of the three variables A,B,C assumes the values $+1$ and -1 with probability $1/2$. If, on the other hand, the twelve entries in Table 1 are all replaced by $1/4$, then three, even independent, random variables A,B,C *can be defined on some common probability space* and each assuming the values $+1$ and -1 with probability $1/2$.

The following question arises immediately. If the twelve entries in Table 1 are replaced by non-negative numbers such that the entries in each row add up to 1, in other words if Table 1 constitutes a 3x4 stochastic matrix, then under which circumstances is it possible to realize the corresponding joint distributions on a common probability space? Bell, unknowingly, addressed this question in the context of quantum mechanics by replacing the twelve entries by numbers that depended on the covariances, resulting from such joint distributions, and that were based on the negative cosines of certain pairs of angles. He then assumed that a joint distribution exists by defining A,B,C as functions of a single random variable Λ as well as certain given settings as indicated in Table 1 with the goal of deducing consequences of "some condition of locality, or of separability of distant systems" [2]. Indeed, Bell tried to show then via his inequality that objective local hidden variables such as Λ can not exist [3] and states: "In a theory in which parameters are added to quantum mechanics to determine the results of individual measurements, without changing the statistical predictions, there must be a mechanism whereby the setting of one measuring device can influence the reading of another instrument, however remote". The basis for this far reaching statement was the fact that for certain combinations of angles, depending on the instrument settings, a contradiction between the predictions of quantum mechanics and his inequality could be obtained. However, as soon as A,B,C are assumed to depend only on the single random variable $\Lambda(\omega), \omega \in \Omega$, then $A = A(\Lambda(\omega)), B = B(\Lambda(\omega))$ and $C = C(\Lambda(\omega))$ are all defined on the same common probability space Ω. Thus in view of Vorob'ev's example the contradiction is obtained even before we get started and without recourse to any inequality. In the same vein, the derivation of the Bell-type inequalities requires no further assumptions and is independent of any additional considerations involving the Einstein separation principle. In the present scenario, the basis for Bell's proof is simply the assumption that A,B,C can simultaneously be measured in the sense that the values that these three random variables assume can be simultaneously registered. Therefore the definition of $A(\Lambda(\omega))$ etc. (denoted by $A(\mathbf{a},\lambda)$ etc. in the first equation of Bell's celebrated paper [3]) contains all the information needed to derive the inequalities that contradict some quantum results. However, according to Vorob'ev's example of Table 1 probabilities arising from certain "closed loops" can not consistently be described by random variables on a common probability space. Hence the contradiction between the Bell inequality and the predictions of quantum mechanics has its roots entirely in purely mathematical reasons. At this point it is of no concern whether or not Λ depends on all or on none of the instrument settings.

To better illustrate the ramifications of this discussion we ask the reader to imagine the following situation. Assume that the Aspect experiment [4] had already been performed and assume that Bell knew about it and also knew Vorob'ev's example of Table 1. Bell wishes now to investigate the possibility of a hidden variable model for the Aspect experiment. He knows that in view of the Vorob'ev example the results of Aspect et al. [4] can not be explained by a model that uses three random variables defined on a common probability space (see also the discussion after Eq.(4)). Therefore he rejects the Ansatz $A = A(\Lambda(\omega)), B = B(\Lambda(\omega))$ and $C = C(\Lambda(\omega))$. No inequalities are used or needed.

However, history proceeded along a different path. The Bell inequality, as well as the more general CHSH [5] inequality, were obtained first and provided the decisive motivation for the Aspect experiment [4]. Λ's independence of the settings, a consequence of the delayed choice of the settings in the Aspect experiment and of Einstein locality, was considered crucial. Bell's Ansatz was considered to be most general and the contradiction of Bell-type inequalities with the data of the Aspect experiment was attributed by Bell to non-localities.

We assume throughout as the basis for our analysis that the Aspect experiment and all its results are valid and that no practical deviations from the ideal embodiment of all experimental procedures, such as detector inefficiencies etc., are of any significance. We believe that ultimately it ought to be the goal to find a physically reasonable mathematical model that can explain the data of the Aspect experiment. We emphasize that our criticism is directed only at some of the previous mathematical models for the Aspect experiment, such as the Bell inequality, the CHSH inequality and the arguments leading to these inequalities.

In many sciences it is a commonly accepted principle that if there are competing theories that can be used to explain certain phenomena then the simplest theory is the chosen one. We believe that a simpler and thus a better explanation of the data of the Aspect experiment can be based on a model that in addition to a source parameter Λ includes time and setting dependent instrument parameters $\Lambda^*_{\mathbf{a}}(t)$ and $\Lambda^{**}_{\mathbf{b}}(t)$. These parameters are Einstein-local and they may be quantum mechanical in nature, in the sense of describing atomistic effects. $\Lambda^*_{\mathbf{a}}(t)$ and $\Lambda^{**}_{\mathbf{b}}(t)$ are permitted to have their own distinctive stochastic behavior and the same is true for all other settings.

We also note that independent of our considerations we believe that the Aspect experiment is crucial for the interpretation of quantum mechanics and particularly for the following important distinction. It may either prove the existence of time and setting dependent equipment parameters or, if their existence can be ruled out for not yet known physical reasons, it may just show what the followers of Bell have deduced all along e.g. non-locality or the lack of validity of counterfactual reasoning as proposed by Peres [6].

In contrast we shall show that other related experiments, particularly those performed by Pan et al. [7] according to the Greenberger-Horne-Zeilinger (GHZ) theory [8], do permit the construction of an objective local model. We shall point out a serious flaw in their theory and, in addition, we will discuss a simple experiment that can be performed on three independent computers and that simulates the same experimental data as the test performed by Pan et al. [7]. This computer simulation directly contradicts their claim that their experiment establishes non-locality without invoking inequalities. We also note that the variations of Hardy, Peres and Mermin on the GHZ theory [9] suffer from similar deficiencies.

CRITIQUE OF PROOFS OF THE CHSH INEQUALITY

Probabilistic proofs

This section is a brief summary of our previous work [10] with the role of the Vorob'ev theorem [1] woven in. First we recall that the Bell inequality is contained in the CHSH inequality as a special case. Therefore we focus our discussion on the CHSH inequality noting that the same arguments apply mutatis mutandis to the Bell inequality. From now on we use the standard set-up and standard notation [10]. Correlated pairs of particles are emitted from a source S_0, and the information they carry is characterized by a random variable Λ. Following Bell, we introduce random variables that describe spin measurements $A = \pm 1$ in station S_1, and $B = \pm 1$ in station S_2. A and B are assumed to be functions of Λ and of the instrument settings that are denoted by three-dimensional unit vectors, usually $\mathbf{a}$ and $\mathbf{d}$ in S_1, and $\mathbf{b}$ and $\mathbf{c}$ in S_2. The instrument settings have a special status in the sense that they are controlled by the experimenters in S_1 and S_2 respectively. The experimenter in S_1 chooses $\mathbf{a}$ or $\mathbf{d}$ with probability $\frac{1}{2}$ and the experimenter in S_2, stochastically independent of the choice in S_1, chooses $\mathbf{b}$ or $\mathbf{c}$ with probability $\frac{1}{2}$.

The standard proofs of the CHSH inequality, as presented in most text-books, proceed as follows. An entity

$$\Gamma := A(\mathbf{a},\Lambda(.)).B(\mathbf{b},\Lambda(.)) + A(\mathbf{a},\Lambda(.)).B(\mathbf{c},\Lambda(.)) + A(\mathbf{d},\Lambda(.)).B(\mathbf{b},\Lambda(.)) - A(\mathbf{d},\Lambda(.)).B(\mathbf{c},\Lambda(.)) \tag{1}$$

is defined. Since $A(\mathbf{a},..)$ and $A(\mathbf{d},..)$ can be factored and since either $B(\mathbf{b},..)+B(\mathbf{c},..) = 0$ or $B(\mathbf{b},..)-B(\mathbf{c},..) = 0$, whereas the other one accordingly equals ± 2 it is then claimed that $\Gamma = \pm 2$. Hence the absolute value $|\Gamma| = +2$, and so integration of this equation with respect to a probability measure and an application of an elementary inequality for integrals yields the CHSH inequality,

$$|E(A(\mathbf{a})B(\mathbf{b})) + E(A(\mathbf{a})B(\mathbf{c})) + E(A(\mathbf{d})B(\mathbf{b})) - E(A(\mathbf{d})B(\mathbf{c}))| \leq 2 \tag{2}$$

Here E stands for the expectation value operator. This proof of the CHSH inequality is based on a mathematical model that does not adequately represent the Aspect experiment. We first argue on the level of elementary calculus. In each run of the Aspect experiment only one of the four products in Γ can be measured, a fact that is generally also appreciated [11], [12]. Γ itself is not measured directly since measurement of the four products requires four incompatible arrangements. It follows that, in particular, the various A's and B's need not be the same, although they are denoted the same. Thus the conclusion that necessarily $\Gamma = \pm 2$ is not correct.

If we try to interpret the above proof of the CHSH inequality as a proof based on measure theoretic probability theory we first need to agree on the underlying sample space. According to the classical basic texts on probability theory by Feller [13] and on measure theory by Halmos [14] for probability theory to apply to real world problems, a one-to-one correspondence between the elements of the sample space and the experiments to be performed must be established first. A random variable is by definition a function (measurable in the sense of measure theory) defined on that sample space, i.e. to each performed experiment there is a well-defined value attached to it, the value that the random variable assumes. Because, as was noted above, the Aspect experiment does not measure Γ itself, without any additional assumptions on the stochastic relationship between the potential hidden variables and the setting vectors, Γ may not be a well-defined function on any sample space, thus may not be a random variable. This is not just a mathematical technicality because the following issue is important. It has been known (see e.g. the critiques by L. Accardi [15] and A. Fine [16]), that the crucial ingredient in the standard proof of the CHSH inequality, as reproduced above, was the assumption that the four random variables $A(\mathbf{a}), A(\mathbf{d}), B(\mathbf{b}), B(\mathbf{c})$ can be defined on the same probability space, thereby equating the various A and B in Γ. At first glance this only seems to contradict the fact that the settings can not be all simultaneously considered which is, in many cases, only a technicality. For example, a fair die can be thrown six times or six fair dice can be thrown once, the resulting statistics is all the same. However, in order to establish that Γ is a random variable, that $\Gamma = \pm 2$ and in turn to take the expectation value $E(\Gamma)$ resulting in the CHSH inequality requires that the joint distributions of the four pairs $[A(\mathbf{a}), B(\mathbf{b})], [A(\mathbf{a}), B(\mathbf{c})], [A(\mathbf{d}), B(\mathbf{b})]$ and $[A(\mathbf{d}), B(\mathbf{c})]$ can be realized as the marginal distributions of the fourfold distribution of [A(**a**), A(**d**), B(**b**), B(**c**)]. For certain special cases this can indeed be established (see proofs at the end of this section). But, in general, this clearly contradicts the conclusion of the Vorob'ev theorem [1] because, picturesquely speaking, the four joint distributions form a closed loop. To be more specific we consider the following table.

TABLE 2. Illustration of the Vorob'ev theorem in terms of the covariances σ.

	$(+1,+1)$	$(+1,-1)$	$(-1,+1)$	$(-1,-1)$
(A(**a**), B(**b**))	$(1+\sigma_{\mathbf{ab}})/4$	$(1-\sigma_{\mathbf{ab}})/4$	$(1-\sigma_{\mathbf{ab}})/4$	$(1+\sigma_{\mathbf{ab}})/4$
(A(**a**), B(**c**))	$(1+\sigma_{\mathbf{ac}})/4$	$(1-\sigma_{\mathbf{ac}})/4$	$(1-\sigma_{\mathbf{ac}})/4$	$(1+\sigma_{\mathbf{ac}})/4$
(A(**d**), B(**b**))	$(1+\sigma_{\mathbf{db}})/4$	$(1-\sigma_{\mathbf{db}})/4$	$(1-\sigma_{\mathbf{db}})/4$	$(1+\sigma_{\mathbf{db}})/4$
(A(**d**), B(**c**))	$(1+\sigma_{\mathbf{dc}})/4$	$(1-\sigma_{\mathbf{dc}})/4$	$(1-\sigma_{\mathbf{dc}})/4$	$(1+\sigma_{\mathbf{dc}})/4$

Here accordingly in each row σ equals the covariance of (A,B), depending on the settings as indicated in $(A(.), B(.))$. Under the assumption that each A and each B assumes the values $+1$ and -1 with probability $\frac{1}{2}$ the covariances σ uniquely determine the entries in Table 2. Quantum mechanics identifies the σ's in each of the four rows as the negative cosines of the angles between the corresponding setting vectors. The Vorob'ev theorem in conjunction with the Kolmogorov existence and consistency theorem states that, in general, it is not possible to realize four such joint distributions as marginals of a fourfold distribution of four random variables defined on a single probability space. To prove this claim directly we choose the σ's in the first three rows to equal $\frac{1}{\sqrt{2}}$ and the σ in the last row to be $-\frac{1}{\sqrt{2}}$. The above claim follows immediately by an argument similar to the one given in the introduction. Indeed from Table 2 we obtain that for each of the four probabilities (corresponding to each choice of $+1$ or -1)

$$P(A(\mathbf{a}) = -1, B(\mathbf{b}) = -1, B(\mathbf{c}) = \pm 1, A(\mathbf{d}) = \pm 1) \leq (1 - \frac{1}{\sqrt{2}})/4 \tag{3}$$

Adding these four probabilities we have

$$P(A(\mathbf{a}) = -1, B(\mathbf{b}) = -1) \leq 1 - \frac{1}{\sqrt{2}} \tag{4}$$

in contradiction to the value assigned by Table 2.

Hence the fact that for some choices of angles between the setting vectors the CHSH inequality is in contradiction with the predictions of quantum mechanics is not a consequence of some mysterious nonlocal physical phenomena, but rather a straightforward consequence of basic mathematics.

Of course, in some particular cases it may be possible to realize these four joint distributions as marginals of a fourfold distribution, for instance if all four σ's equal zero, i.e. all 16 entries equal $\frac{1}{4}$, i.e. if the four random variables are pairwise independent.

The following facts are now evident. In view of the Vorob'ev theorem and example, *neither the Bell inequality nor the CHSH inequality provide a conclusive tool to decide whether or not an objective local model of any particular experiment can be established*. What is really needed is a direct check of whether or not the relevant random variables can be defined on the same probability space. For the example of Table 1 the Bell inequality,

$$|E(AB) - E(AC)| \leq 1 - E(BC) \tag{5}$$

corresponding to the three angles $45°, 45°, 90°$ resulting in the covariances $\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0$ is satisfied and yet, as we have seen, A, B, C can not be defined on the same probability space. Thus for this case Bell's inequality is fulfilled yet Bell's Ansatz still needs to be rejected.

For the mathematically inclined reader we offer the following comments as to what is precisely going on here. Given three angles or the corresponding unit vectors at least three Bell inequalities can be written down. The one as displayed above and two more obtained by cyclic permutation. Now it is not difficult to see that if the direct approach shows that the three pair distributions cannot be obtained from three random variables defined on the same probability space then one of these Bell inequalities will be violated. Hence the direct approach is equivalent with checking all Bell inequalities, because as the above example shows, checking only one of the inequalities may not be enough to reach the desired conclusion.

On the other hand, for a very small class of parameters (e.g. source parameters [10]) the conclusion of the Bell theorem that is drawn from locality conditions remains valid. This can be proved in several ways. In [10] we have presented proofs based on what we called a reordering argument that work in some special situations. For the benefit of the reader we will reproduce this argument, based on elementary statistics, in the following section. A second way to establish the Bell theorem in this special situation is in essence the probabilistic counterpart of [10] and is as follows. We assume that the hidden variable consists only of a source parameter $\Lambda = \Lambda(\omega)$ that is stochastically independent of the setting vectors, considered as random variables $X(\omega^*)$ assuming one of the two possibilities **a** or **d** in S_1, and $Y(\omega^{**})$ assuming the values **b** or **c** in S_2, with probability $\frac{1}{2}$ each. Because of the hypothesis of the stochastic independence we can assume that there is a common product probability space, say $\Omega \mathrm{x} \Omega^* \mathrm{x} \Omega^{**}$ on which Λ, X and Y are well-defined. Since A and B are assumed to be functions only of x and λ and only of y and λ, respectively, $A(X,\Lambda)$ and $B(Y,\Lambda)$ are random variables defined on the same probability space. Hence applying Fubini's theorem we conclude that $A(\mathbf{a},\Lambda(\omega)), A(\mathbf{d},\Lambda(\omega)), B(\mathbf{b},\Lambda(\omega))$ and $B(\mathbf{c},\Lambda(\omega))$ are also random variables that are all defined on the same probability space. Thus in this special case the entity Γ turns out to be a random variable and hence in this special case the proof of the CHSH inequality is correct. A third way, and at the same time the most efficient one to establish the Bell theorem in this special situation is based on the above example that shows that the four joint pair distributions in Table 2 cannot be obtained as the marginals of four random variables $A(\mathbf{a}), A(\mathbf{d}), B(\mathbf{b})$ and $B(\mathbf{c})$. Just above, in connection with Fubini's theorem we remarked that substitution of Λ would entail the opposite statement. Hence such parameters Λ can be ruled out.

Another, more general instance where the Bell theorem remains valid is when in addition to the source parameter Λ instrument parameters Λ^* in S_1 and Λ^{**} in S_2 are considered such that these three parameters are stochastically independent. The same type of arguments will work in this special case, too. Hence, these special classes of hidden variables can be ruled out, that is for these classes the Bell theorem is correct.

Of course, the validity of the Bell theorem in these special cases does not imply that the Bell theorem is correct in general, that is that all classes of hidden variables that can be reasonably considered of being Einstein-local can be ruled out. For instance the above type of arguments no longer work for what we call the extended parameter space, that consists of a source parameter Λ, and time and setting dependent instrument parameters $\Lambda^*_{\mathbf{a}}(t)$ and $\Lambda^{**}_{\mathbf{b}}(t)$ in S_1 and S_2, respectively. Because of the special role of time it is admissible for the instrument parameters to be correlated by time without violating the principle of Einstein-locality.

Proofs based on elementary statistics

Some standard text books consider the data accumulated by sampling the above Γ. This is just the equivalent statistical realization of the probability model considered in the previous section. However, in general, Γ cannot be sampled because, in general, Γ is not a random variable. A reordering argument works for the special situations considered above. As in the previous section let us assume that Λ and the setting vectors X and Y are stochastically independent and that X assumes the vector values **a** and **d** and Y assumes the vector values **b** and **c** with probability $1/2$ each. To avoid dealing with ε's, let us assume in addition that Λ assumes only finitely many values λ_s with positive probability $p_s, s = 1,2,...,M$. If a large number N of runs of the Aspect experiment is performed then for each $s = 1,2,...,M$ we expect the parameter value λ_s to occur approximately $N.p_s$ times. Since by independence each pair of setting vectors (**a**,**b**), (**a**,**c**), (**d**,**b**) and (**d**,**c**) occurs about $\frac{1}{4}$ of the times each λ_s occurs, at the end of the day we can reorder and rearrange the data points in approximately $\frac{1}{4}N.p_s$ rows all reporting to the same λ_s. Because now the λ_s are the same in each row the corresponding values γ are indeed ± 2. Taking averages, denoted by $< >$ and discarding rows that are possibly incomplete we obtain the CHSH inequality in the form

$$| < a_j.b_j > + < a_j.c_j > + < d_j.b_j > - < d_j.c_j > | \leq 2 \tag{6}$$

Note that the lower case $a, b, ...$ denote the outcomes corresponding to the random variables with the corresponding settings but not the settings themselves which are always boldfaced. This is in agreement with the standard notation [11].

Again this line of argument fails for our extended parameter space.

Proofs based on sampling tables

There are several arguments in the literature linking the validity of the Bell theorem with sampling some kind of table and often without any reference to the issue of hidden variables. In contrast to the above situations in these arguments the emphasis is placed solely on the set of potentially measured data. Most of these arguments are deficient. Below we shall comment only on one of these tables. But first, in analogy to Table 2, we describe the sampling procedure that mimics the Aspect experiment, reviewing again the logical issues that need to be considered when setting up a proper mathematical model. In the following table the 16 possible outcomes of the measurement pairs in the Aspect experiment are represented as follows.

TABLE 3. 16 possible outcomes of a Aspect experiment run

$a(+).b(+)$	$a(+).b(-)$	$a(-).b(+)$	$a(-).b(-)$
$a(+).c(+)$	$a(+).c(-)$	$a(-).c(+)$	$a(-).c(-)$
$d(+).b(+)$	$d(+).b(-)$	$d(-).b(+)$	$d(-).b(-)$
$d(+).c(+)$	$d(+).c(-)$	$d(-).c(+)$	$d(-).c(-)$

Within each row a particular product is chosen with a certain fixed probability such that the probabilities for each row add up to 1. In other words the corresponding table of these probabilities constitutes a 4x4 stochastic matrix. The example given in Table 2 represents the prediction by quantum mechanics. Then each row in Table 3 is sampled according to the probability distribution (for instance, of Table 2) corresponding to that row and the average corresponding to the outcomes of that row is calculated. Then the average of the fourth row is subtracted from the sum of the averages of the first three rows. Call the resulting quantity $<\gamma>$. The CHSH inequality is equivalent to the statement that no matter how the probabilities in the corresponding 4x4 stochastic matrix are chosen, the resulting $<\gamma>$ never exceeds 2 in absolute value. However, it is easy to give examples of 4x4 stochastic matrices with the property that if we sample Table 3 according to the probabilities of such a table then $<\gamma>$ can be certainly bigger than 2, even as big as 4. If, in addition, we also mimic the delayed choice provision of the Aspect experiment then "after emission of a pair of particles" a row of the above table gets chosen with probability $1/4$ and then from this chosen row a sample of size 1 is taken according to the corresponding probability distribution. After a series of such "emissions" $<\gamma>$ is then calculated accordingly. The Aspect experiment shows that the claim that $<\gamma>$ never exceeds 2 in absolute value is false for some 4x4 stochastic matrices and the Vorob'ev theorem provides the mathematical rationale for this fact. Seen from this vantage point, the delayed choice provision in the Aspect experiment is of no consequence.

Another demonstration of the conclusion of the Vorob'ev theorem is given unwittingly in Table 6-1, on page 167 of the text-book by Peres [11]. There Peres gives lower bounds on the covariances of certain double sequences $(a_j,b_j),(b_j,c_j),(c_j,d_j)$ and (d_j,a_j) and then wonders why, at the end, the covariance of the double sequence (d_j,a_j) satisfies two conflicting inequalities. Because, as observed above, the joint distribution of a pair of random variables, each assuming with probability $1/2$ the values +1 and -1 only, can be expressed in terms of their covariance, and since the empirical distributions of the four double sequences, again expressed picturesquely, form a closed loop the Vorob'ev theorem says that, in general, it is not possible to find a consistent joint distribution of four random variables yielding the above imagined data set on which the argument of Peres is based on.

THE GHZ APPROACH

A decisive and penetrating analysis of the GHZ approach has been given by Khrennikov [17]. Hence we shall review here only a few of the basic ingredients needed below to describe the computer experiment. In the paper "Bell's theorem without inequalities", by Greenberger, Horne, Shimony, Zeilinger [18] the implicit assumption that the λ that occurs in all the relations must be the same is clearly unfounded. The results of the actual experiments are reported in [7]. For the photons $i=1,2,3$ the authors introduce "elements of reality X_i with values ± 1 for $H'(V')$ polarizations and Y_i with values ± 1 for R(L)" polarizations. The authors claim that the elements of reality X_i and Y_i satisfy the relation

$$Y_1.Y_2.X_3=-1, \qquad Y_1.X_2.Y_3=-1, \qquad X_1.Y_2.Y_3=-1 \tag{7}$$

Invoking counterfactual reasoning, the authors conclude that

$$Y_i.Y_i = +1 \tag{8}$$

and thus by Eq.(7) that

$$X_1.X_2.X_3 = -1 \tag{9}$$

Counterfactual reasoning by itself is, in our opinion, not objectionable. One certainly can argue that, if I had measured with different settings and if I had the same photon(s) then I would have obtained...In mathematical terms this just means that for a given ω, representing a given experiment, the random variables X_i and Y_i may assume one of the values $X_i(\omega)$ and $Y_i(\omega)$, respectively, where $i = 1,2,3$. However, this fact does not permit the conclusion that in Eq.(7) the two Y_1 necessarily assume the same values, i.e. they are the same random variables, because they are definitely obtained in two distinctly different experiments. A similar statement holds for Y_2 and Y_3. Thus subject to this interpretation Eq.(8) is false as it stands.

Proceeding now with the discussion of the experiments of Pan et al. [7] we note that Figure 3 of their paper depicts the histograms for the actual *yyx*, *yxy*, and *xyy* experiments. The error rate for each of these three experiments is given with 0.15 ± 0.02. This translates into an error rate for the *xxx* result extracted from Eq.(7) of 0.45 ± 0.035 which is close to 50%. Figure 4 of their paper shows the *xxx* results measured directly in a separate experiment. With success rate 0.87 ± 0.04, it demonstrates that the product in Eq.(9) equals +1. This contradiction is statistically highly significant and is the basis for their claim that, as a consequence of their test, for the three-photon entanglement the "quantum physical predictions are mutually contradictory with expectations based on local realism."

We shall show now that their claim of non-locality is false by providing an example that can be simulated on three independent computers. Let

$$r_k(t) = sign[sin(2^k \pi t)] \text{ for } t > 0 \tag{10}$$

denote the k-th Rademacher function. Note that r_k has period $2^{-(k-1)}$. The following table can serve as a basis for this simulation

TABLE 4. Computer simulation of the experiment of Pan et al. [7]

	$yyx, t_0 < t < t_1$	$yxy, t_2 < t < t_3$	$xyy, t_4 < t < t_5$	$xxx, t_6 < t < t_7$
Comp1	$Y_1 = -r_1$	$Y_1 = -r_1$	$X_1 = r_2.r_3$	$X_1 = r_2.r_3$
Comp2	$Y_2 = r_2$	$X_2 = r_1.r_3$	$Y_2 = r_2$	$X_2 = r_1.r_3$
Comp3	$X_3 = r_1.r_2$	$Y_3 = r_3$	$Y_3 = -r_3$	$X_3 = r_1.r_2$

Here $t_1 - t_0$ is the length of time the *yyx* experiment is running, $t_2 - t_1$ is the length of time it takes the experimenters to switch the experimental set-up from an *yyx* experiment to an *yxy* experiment. $t_3 - t_2$ is the length of time the *yxy* experiment is running and $t_4 - t_3$ again the time to switch and so forth as described in Table 4. Each of the three equations in Eq.(7) holds on the entire time interval where they are defined. Moreover, we have

$$X_1.X_2.X_3 = +1 \tag{11}$$

instead of Eq.(9), if we mimic at a later time the *xxx* experiment according to the last column in Table 4. Furthermore, each X and each Y equals +1 or -1 half of the time. The essential point here is, of course, that for given equipment settings, e.g. *yxy*, we can assume that equipment parameters are such that Y may be described by a certain Rademacher function, e.g. $Y_3 = r_3$, while for the other *xyy* we may have $Y_3 = -r_3$. Here we have made use of the fact that in the actual experiment the settings e.g. *yyx* are set and used for a longer period of time so that a mutual report can be established by sub-light velocities between the measurement stations as to which overall setting (*yxy* or *xyy* etc.) is used. For a given setting, the outcomes of the various experiments are, of course, only "known" at a given detector, not at the others. Only the choice of measurement time, which is random, determines the outcomes together with the Rademacher functions that are characteristic for a given setting. Of course the three Rademacher functions in Table 4 can be replaced by three other Rademacher functions with arbitrarily large but different subscripts if faster fluctuation between +1 and -1 is desired.

CONCLUSIONS

We have given a summary of the reasons why most of the proofs leading to the Bell inequality and to the CHSH inequality are deficient. In particular we have shown that in view of the theorem of Vorob'ev the possibility of agreement of these inequalities with the Aspect experiment is immediately lost as soon as A, B are assumed to be the functions defined by Bell: his inequality follows independently of any physics or locality conditions. We have shown that time and setting dependent instrument parameters that are Einstein-local *need not satisfy* Bell-type inequalities. In fact, none of the known arguments leading to the CHSH inequality can accommodate these parameters. Also, we have presented a method more efficient than the Bell and the CHSH inequality that can help to weed out specific classes of hidden variables. This method does not rely on inequalities, but rather on a simple determination whether a given set of joint pair distributions can be realized as the marginals of the joint distribution of random variables defined on the same probability space. As mentioned in [10], we do not have a proof that in reality, not just mathematically, these setting and time dependent parameters do exist for the Aspect experiment. Such a proof would be established indirectly if, for instance, one would be able to play the Bell game on two (stochastically and/or functionally) independent computers with the same clock time. Given no further information, we see this as a very difficult problem. However, playing a Bell-type game for the GHZ approach is relatively easy as we have demonstrated in the last section. For GHZ type of experiments we do have an existence proof for setting and time dependent instrument parameters because of the possibility of the computer experiment. We believe that it is only a matter of time that the same will be found for the Aspect experiment. Using the (perhaps somewhat unreasonable) physical assumption of history dependent instrument outcomes, we have found it already [19].

ACKNOWLEDGEMENTS

The authors would like to thank Andrei Khrennikov and Louis Marchildon for valuable discussions. Support of the Office of Naval Research (N00014-98-1-0604) is gratefully acknowledged.

REFERENCES

1. N. N. Vorob'ev, *Theory of Probability and its Applications* **VII**, 147-162 (1962).
2. J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press, Cambridge, U.K., 1993, p 11.
3. J. S. Bell, *Physics* **1**, 195-200 (1964).
4. A. Aspect, J. Dalibard and G. Roger, *Phys. Rev. Letters* **49**, 1804-1807 (1982).
5. J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt *Phys. Rev. Letters* **23**, 880-883 (1969).
6. A. Peres *Am. J. Phys.* **46**, 745-747 (1978).
7. Jian-Wei Pan, D Bouwmeester, M. Daniell, H. Weinfurter, H. and A. Zeilinger, *Letters to Nature* **403**, 515-519 (2000).
8. D. M. Greenberger, M. A. Horne and A. Zeilinger, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, edited by M. Kafatos, Kluewer Academic Publishers, Dordrecht, 1989, pp 73-76.
9. A. Peres, '*Quantum Theory: Concepts and Methods*, Kluewer Academic Publishers, Dordrecht, 1995, pp 151-153.
10. K. Hess and W. Philipp, *Proceedings of the National Academy of Sciences (USA)* **101** 1799-1805 (2004).
11. A. Peres, '*Quantum Theory: Concepts and Methods*, Kluewer Academic Publishers, Dordrecht, 1995.
12. A. J. Leggett, *The Problems of Physics*, Oxford University Press, 1987, pp 144-172.
13. W. Feller, *An Introduction to Probability Theory and its Applications* Vol 1, Wiley Series in Probability and Mathematical Statistics, New York, 1968, pp 1-9.
14. P. R. Halmos, *Measure Theory*, Van Nostrand Co., Priceton, N. J., 1950, p 188.
15. L. Accardi, *Phys. Rev.* **77**, 169-192 (1981).
16. Fine, A. *J. Math. Phys.* **23**, 1306-1310 (1982).
17. Andrei Khrennikov, *Physics Letters A* **278**, 2001, pp 307-314.
18. D. M. Greenberger, M. Horne, M.A. Shimony and A. Zeilinger, *Am. J. Phys* **58**, 1131 -1143 (1990).
19. K. Hess and W. Philipp, *Proceedings of the National Academy of Sciences (USA)***98**, 14228-14233 (2001). See also: W. Philipp and K. Hess, Proc. Conf. Foundations of Probability and Physics -2, edited by A. Khrennikov, Vaxjo, Sweden, 2002, pp 469-488.