How often are Bell's inequality premises violated?

1980 J. Phys. A: Math. Gen. 13 3187

(http://iopscience.iop.org/0305-4470/13/10/015)

J. Phys. A: Math. Gen. 13 (1980) 3187–3194. Printed in Great Britain

How often are Bell's inequality premises violated?

Susan J Feingold†§ and Asher Peres‡

† Technion Research and Development Foundation Ltd, Haifa, Israel
‡ Department of Physics, Technion-Israel Institute of Technology, Haifa, Israel

Received 13 February 1980, in final form 21 March 1980

Abstract. When the polarisations of two photons are correlated, the result of a measurement of a polarisation component of one of the photons is not independent of the *choice* of the component measured on the *second* photon (even if the two measuring instruments are very distant). In this paper, we compute the probability of a change in the readings of the first instrument, associated with a given change of orientation of the second instrument. Under optimal conditions, we show that in at least 41% of cases ($\sqrt{2}-1$) the readings of one of the instruments cannot be independent of a reorientation of the other one.

1. Introduction and glossary

In general, quantum theory gives us only statistical predictions for the outcomes of measurements. It is possible to explain (or at least to simulate) these statistical laws by deterministic theories with additional 'hidden' variables (Bell 1966, Bohm and Bub 1966) contrived in such a way that an average over these variables gives results identical to the quantum mechanical average. However, the following remarkable theorem was proved long ago by Bell (1964). If two systems have correlated properties because of their past history, it is impossible to separate their hidden variables in two subsets, such that the outcomes of measurements performed on each system depend only on one of these subsets.

It was later shown that the same property holds for systems correlated by their *future* history (Costa de Beauregard 1979a) and that it can be formulated in purely macroscopic terms without any reference to hidden variables (Stapp 1971, Eberhard 1977, Peres 1978). Namely, consider the outcomes of all possible pairs of measurements which can be devised for the two correlated systems, even though the actual performance of one such pair of measurements precludes the simultaneous performance of any other pair.† Moreover, assume that the readings of each instrument are not affected by the experimental set-up of the other instrument. Then the outcomes of *all* these measurements (whether actually performed or not) *cannot satisfy correlations as strong as those requested by quantum theory.* They can only satisfy the *weaker* classical correlations (Peres 1978).

This astonishing result seems to imply the existence of an instantaneous 'influence' conveyed over arbitrarily large distances‡ (Stapp 1977, Costa de Beauregard 1978,

§ Now at Elbit Computers Ltd, Haifa, Israel.
† E.g., in the simpler case of a *single* particle, we can devise measurements of *q or p*, but the actual performance of one of these measurements precludes the performance of the other.
‡ This effect was foreseen by Einstein (1949) who used the word 'telepathically'.

Slaby 1978). The purpose of our work is to calculate the probability of a change in the readings of one instrument, in relation to alternative set-ups of the *other* instrument. Our calculations are model independent. In particular they do not involve the use of hidden variables.

Consider as an example the experimental set-up proposed by Aspect (1976) and schematically described in figure 1. It consists actually of four different experiments, of which only one is actually performed for each photon pair. We shall use the following terminology: each pair of angles of the analysers ($\alpha\beta$, $\alpha\delta$, $\gamma\beta$ or $\gamma\delta$) corresponds to an *experiment.* The set of possible observations of a pair of orthogonally polarised photons will be called a *run* (Davidon 1977). Therefore each run involves four different experiments. Only one of these is actually performed and we can try to guess (or to compute, if we have a suitable theory) the results of the three other experiments.

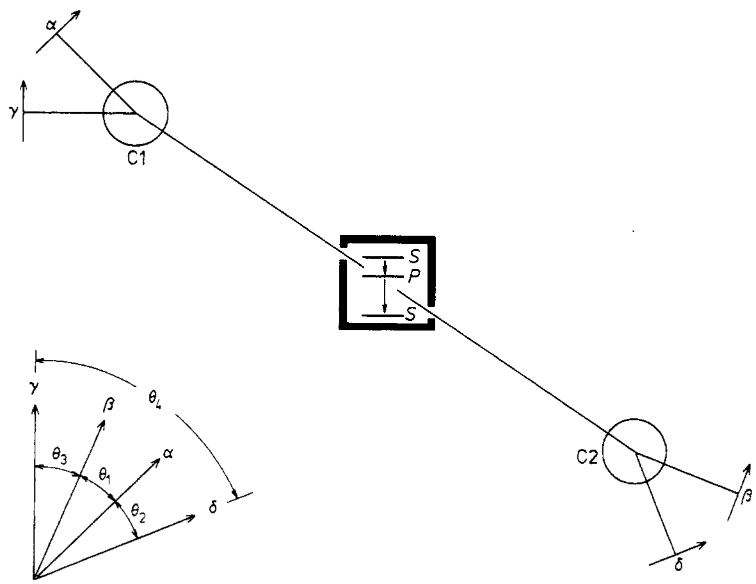

Figure 1. Two photons with opposite polarisations are emitted in an S–P–S cascade. They pass through random optical commutators C1 and C2 toward polarisation analysers oriented along the directions α, β, γ and δ. There are four different pairs of angles ($\alpha\beta$, $\alpha\delta$, $\gamma\beta$ and $\gamma\delta$) and therefore four different experiments which are simultaneously considered. Of course, only one of these experiments is actually performed for each photon pair.

The diagram in the left corner shows the set of angles (three times 22·5°) giving the maximum violation of Bell's inequality.

Let A, B, C or D be defined as $+1$ if a photon has passed through one of the analysers at angles α, β, γ or δ respectively, -1 if it has been rejected by that analyser. The results (observed or computed) of a series of runs can be represented by table 1. Because of rotational symmetry, their average values must be

$$\langle A\rangle = \langle B\rangle = \langle C\rangle = \langle D\rangle = 0, \tag{1}$$

which can be trivially satisfied. Moreover, since the photons have orthogonal polarisations, there are *correlations* between the readings of the instruments. According to

Table 1. Outcomes (real or imagined) of a series of runs, assuming that each reading of analyser α or γ is independent of whether the other analyser is β or δ, and vice versa. Here and in the following tables, + and − mean ± 1, for brevity.

Run no.	Readings of instruments A	B	C	D
1	+	−	−	+
2	+	+	−	−
3	−	+	+	−
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
N	.	.	.	.

quantum theory we have, for each of the four possible experiments (Clauser and Shimony 1978):

$$\begin{aligned} \langle AB\rangle &\equiv Q_1 = -\cos 2\theta_1, \\ \langle AD\rangle &\equiv Q_2 = -\cos 2\theta_2, \\ \langle CB\rangle &\equiv Q_3 = -\cos 2\theta_3, \\ \langle CD\rangle &\equiv Q_4 = -\cos 2\theta_4, \end{aligned} \tag{2}$$

where θ denotes the relative angle, i.e., $\theta_1 = \beta - \alpha$, $\theta_2 = \alpha - \delta$, $\theta_3 = \gamma - \beta$ and $\theta_4 = \gamma - \delta$.

In fact, it is not necessary to make use of quantum theory: each one of the above relations may simply be considered as an empirical result for the corresponding experiment (Freedman and Clauser 1972, Clauser 1976, Fry and Thompson 1976). Therefore the *calculations of the present paper are independent of the validity of quantum theory* (as long as the correlations Q_i depend only on the relative angles of the analysers).

The outline of this paper is as follows. In § 2, we show that table 1 is incompatible with the correlations (2). In other words, it is impossible to assume that the readings of an analyser at angle α or γ are independent of whether the *other* angle is β or δ, and vice versa. We must therefore specify whether we are measuring A_β (first experiment) or A_δ (second experiment) and likewise for the other variables. Consequently, table 1 must be replaced by a more complicated table with *eight* + or − signs for each run. This point is discussed in § 3, where the runs are divided in various *classes*: a *run of class n* (where $n = 0, \ldots, 4$) is one where n of the following equations are *violated*:

$$\begin{aligned} A_\beta &= A_\delta, \qquad & C_\beta &= C_\delta, \\ B_\alpha &= B_\gamma, \qquad & D_\alpha &= D_\gamma. \end{aligned} \tag{3}$$

The runs of class 0 are those which satisfy intuitive separability: two distant instruments do not affect each other's readings. In runs of class 1, the instantaneous 'influence' has altered one of the readings, in runs of class 2, two of them, etc.

Our problem is to find what is the minimum percentage of runs of class 1 (or higher) so that the correlations (2) may be satisfied. This problem is solved in § 4. It is found that although this minimum is unique, the solution corresponding to it is not. In particular, we can arbitrarily impose that *any three* of equations (3) be satisfied, without affecting the total amount of nonlocal 'influence' needed to satisfy equations (2).

Finally, § 5 discusses the meaning of our results in a relativistic theory, where the notion of simultaneity is no longer valid.

2. Bell's inequality

For each run in table 1, there are $2^4 = 16$ different possible outcomes. In order to simplify the calculation of the 16 relative frequencies of the various types of runs, we note that equations (2) remain invariant under the substitution $A \to AD$, $B \to BD$, $C \to CD$ and $D \to D^2 = 1$. We can therefore assume without loss of generality that $D = 1$. This amounts to keeping unchanged the subset of runs having $D = +1$, and changing all signs in the other subset. Of course with that convention, we no longer have to satisfy equations (1).

We thus have only 8 different outcomes, the relative frequencies of which are given in table 2. We readily obtain, from equations (2)

$$\begin{aligned} Q_1 &= (p-v)-(q-w)-(r-u)-(s-t), \\ Q_2 &= -(p-v)+(q-w)-(r-u)-(s-t), \\ Q_3 &= -(p-v)-(q-w)+(r-u)-(s-t), \\ Q_4 &= (p-v)+(q-w)+(r-u)-(s-t), \end{aligned} \tag{4}$$

which we can solve as

$$\begin{aligned} p-v &= (Q_1-Q_2-Q_3+Q_4)/4, \\ q-w &= (-Q_1+Q_2-Q_3+Q_4)/4, \\ r-u &= (-Q_1-Q_2+Q_3+Q_4)/4, \\ s-t &= (-Q_1-Q_2-Q_3-Q_4)/4. \end{aligned} \tag{5}$$

We are now faced with the following consistency problem: each one of the relative frequencies $p, \ldots, w$ must be non-negative and their sum must be equal to one.

Since $-1 \leq Q_i \leq 1$, each one of the expressions (5) also lies between -1 and 1. Therefore the positivity requirement alone for $p, \ldots, w$ never causes any difficulty. However, the sum rule may sometimes be impossible to satisfy. E.g. consider the case

Table 2. Relative frequencies of outcomes A, B and C, assuming that $D = 1$ and that the readings of each instrument are independent of the orientation of the other instrument.

Relative frequency	Readings of instruments A	B	C	D
p	−	−	+	+
q	+	−	+	+
r	−	+	+	+
s	−	+	−	+
t	+	+	+	+
u	+	+	−	+
v	+	−	−	+
w	−	−	−	+

represented in figure 1, where the angles $2\theta_1$, $2\theta_2$ and $2\theta_3$ are acute, while $2\theta_4$ is obtuse, so that Q_1, Q_2 and Q_3 are negative, but Q_4 is positive. Since

$$(Q_4 - Q_1 - Q_2 - Q_3)/2 = (p+q+r+s) - (t+u+v+w), \tag{6}$$

the LHS of this expression should be less than one (this is Bell's inequality).

Now according to the correlations (2)—which can be either derived from quantum theory or simply measured experimentally—the LHS of (6) may exceed 1. Its maximum value is $\sqrt{2}$, when $\theta_1 = \theta_2 = \theta_3 = \pi/8$ and $\theta_4 = 3\pi/8$. This is due to the fact that the cosine correlation is 'too strong' (Herbert 1975). The weaker classical correlations (Peres 1978), namely

$$Q_{\text{classical}} = -1 + 4\theta/\pi, \tag{7}$$

would yield exactly 1 for the LHS of (6). In that case, we would have $t = u = v = w = 0$ and p, q, r and s would be uniquely given by equations (5).

Obviously, the logical flaw in the preceding calculations was to assume that the readings of each instrument are not affected by the orientation of the other one. We shall call this assumption 'intuitive separability'. Actually, when we measure A, we should specify whether this is done as part of the first experiment (other analyser along direction β) or of the second experiment (other analyser along direction δ). Denoting the results as A_β and A_δ respectively, with similar notations for the other analysers, the intuitive separability conditions are given by equations (3). These conditions were hitherto assumed as self-evident. We now see that at least one of them must be violated. The question to which we address ourselves in this paper is: how often must this happen?

3. Construction of a consistent set of runs

If equations (3) are no longer taken for granted, each column of readings in table 1 must be replaced by a pair of columns, e.g. column A becomes columns A_β and A_δ. There are now $2^8 = 256$ types of runs, which can be subdivided into various *classes*, depending on how many of equations (3) are violated. There are 16 different runs of class 0 (none of these equations is violated), 64 of class 1 (one is violated), 96 of class 2, 64 of class 3 and 16 of class 4.

In this section we shall assume for simplicity that we have only runs of classes 0 and 1. A consistent set of such runs is constructed in table 3. The generalisation to arbitrary runs will be discussed in the next section, where we show that this set actually minimises the violation of the separability equations (3).

To construct table 3, we have assumed that $D_\gamma = 1$ (this is similar to the assumption $D = 1$ in the construction of table 2 and does not restrict the generality of the results). The first four rows of table 3 satisfy the separability conditions (3) and are actually the same as the first four rows of table 2. On the other hand, *none* of the rows of table 3 corresponds to the last four rows of table 2. Our reason for this choice can be seen in equation (6): as the difficulty was that its LHS was 'too large', we therefore make the RHS as large as possible by setting $t = u = v = w = 0$, as in classical correlations.

Even with this choice, the cosine correlations are *too strong* (their absolute values are *too large*) to satisfy all the equations (2) with the first four rows of table 3. We therefore tentatively add a fifth row, with runs such that $A_\beta \neq A_\delta$, but satisfying the

Table 3. A set of instrument readings (real or imagined) consistent with the quantum correlations, equations (2). These readings satisfy at least three of the separability conditions, equations (3). It will be shown in § 4 that this is the 'best' set, minimising the violation of equations (3).

Relative frequencies of runs	Readings of instruments							
	Experiment 1		Experiment 2		Experiment 3		Experiment 4	
	A_β	B_α	A_δ	D_α	C_β	B_γ	C_δ	D_γ
$(1+Q_1)/2$	−	−	−	+	+	−	+	+
$(1+Q_2)/2$	+	−	+	+	+	−	+	+
$(1+Q_3)/2$	−	+	−	+	+	+	+	+
$(1-Q_4)/2$	−	+	−	+	−	+	−	+
†	+	−	−	+	+	−	+	+
†	−	+	−	+	+	−	+	+
†	−	+	−	+	−	+	+	+
†	+	−	+	−	+	−	+	+

† The *sum* of the relative frequencies of the last four rows is $\frac{1}{2}(Q_4 - Q_1 - Q_2 - Q_3) - 1$.

three other equations (3). Since we have to *increase* the absolute values of the correlations, we choose

$$
\begin{aligned}
1 &= C_\delta && \text{to increase } Q_4 = \langle C_\delta D_\gamma \rangle, \\
&= C_\beta && \text{by virtue of (3)}, \\
&= -B_\gamma && \text{to increase } |Q_3| = -\langle C_\beta B_\gamma \rangle, \\
&= -B_\alpha && \text{by virtue of (3)}, \\
&= A_\beta && \text{to increase } |Q_1| = -\langle A_\beta B_\alpha \rangle, \\
&= -A_\delta && \text{because the first of equations (3) is violated}, \\
&= D_\alpha && \text{to increase } |Q_2| = -\langle A_\delta D_\alpha \rangle, \\
&= D_\gamma && \text{by virtue of (3)}.
\end{aligned}
$$

The last line coincides with our initial assumption and shows the consistency of these relations.

We have thereby constructed the fifth row of table 3. Assuming that the following rows do not appear, we can now easily find the relative frequencies of these five types of runs, so that the quantum correlations (2) are satisfied. For the first three types, the relative frequencies turn out to be

$$(1+Q_i)/2 = \sin^2 \theta_i, \tag{8}$$

for the fourth one,

$$(1-Q_4)/2 = \cos^2 \theta_4, \tag{9}$$

and for the last one, which violates separability, it is

$$\tfrac{1}{2}(Q_4 - Q_1 - Q_2 - Q_3) - 1 = \sin^2 \theta_4 - \sin^2 \theta_1 - \sin^2 \theta_2 - \sin^2 \theta_3. \tag{10}$$

Note that this is exactly the amount by which Bell's inequality is violated. It can be as large as $\sqrt{2} - 1 \simeq 41\%$ of the runs.

Exactly the same results are obtained if instead of $A_\beta \neq A_\delta$, we assume that any other of equations (3) is violated. Therefore in the final form of table 3, we cannot specify the individual relative frequencies of the various runs of class 1, but only their *total* relative frequency, equation (10).

4. Effect of higher class runs

In the preceding section we have shown that if we restrict ourselves to runs of types 0 and 1 (if at most one of equations (3) is violated) a consistent set of relative frequencies is given by table 3. It would be desirable to have a formal proof that this is indeed the solution which minimises the violation of equations (3). A rigorous mathematical proof is certainly possible (there are only 128 variables) but undoubtedly very cumbersome. We shall therefore proceed in the same heuristic way which already led us to eliminate runs such as those of the last four lines of table 2.

Since the violation of equations (3) is needed to *increase* the absolute values of the correlations, we can write (starting as before from D_γ)

$$D_\gamma = C_\delta \stackrel{?}{=} C_\beta \neq B_\gamma \stackrel{?}{=} B_\alpha \neq A_\beta \stackrel{?}{=} A_\delta \neq D_\alpha \stackrel{?}{=} D_\gamma, \tag{11}$$

where $\stackrel{?}{=}$ means either $=$ or $\neq$. Consistency of the first and last terms implies that either one or three of the $\stackrel{?}{=}$ signs means $\neq$. In other words, only runs of classes 1 and 3 can help to increase *all* the $|Q_i|$. On the other hand, runs of even class would decrease at least one of the $|Q_i|$. We would then need additional runs of odd class to compensate their effect.

The four possible runs of class 3 which satisfy equation (11) are given in table 4. It is seen that they improve the $|Q_i|$ in exactly the same amount as the runs of class 1 (see table 3). They only involve more nonlocality.

It is therefore plausible that table 3 is indeed the 'best' solution of our problem. We see that the nonlocal effect foreseen by Einstein (1949) and formally proved by Bell (1964) is quite large: if $\theta_1 = \theta_2 = \theta_3 = \pi/8$, more than 41% of runs *cannot* satisfy equations (3). They must violate the intuitive requirement that the readings of one instrument are not affected by a change of orientation of the other, distant instrument.

Table 4. Runs of class 3 which affect all the correlations in the desirable direction.

Remaining separability condition	Readings of instruments: Experiment 1 A_β	Experiment 1 B_α	Experiment 2 A_δ	Experiment 2 D_α	Experiment 3 C_β	Experiment 3 B_γ	Experiment 4 C_δ	Experiment 4 D_γ
$A_\beta = A_\delta$	+	−	+	−	−	+	+	+
$B_\alpha = B_\gamma$	−	+	+	−	−	+	+	+
$C_\beta = C_\delta$	−	+	+	−	+	−	+	+
$D_\alpha = D_\gamma$	+	−	−	+	−	+	+	+

5. Remarks on causality

Up to now, we have tacitly assumed that the two halves of each experiment are simultaneous. However, this property was never used explicitly. It is quite possible to arrange the experimental set-up (figure 1) so that its $C_1\alpha\gamma$ part is in the past light cone of $C_2\beta\delta$ (Costa de Beauregard 1979b). In that case, it is reasonable to expect that $A_\beta = A_\delta$ and $C_\beta = C_\delta$, since these readings can be irreversibly recorded *before* the optical commutator C_2 has to make its choice between β and δ. However, we have already seen that our intuition may fail. It is therefore gratifying that the solution of our problem (table 3) is not unique. We can arbitrarily impose $A_\beta = A_\alpha$ and $C_\beta = C_\delta$ (or $B_\alpha = B_\gamma$ and $D_\alpha = D_\gamma$) without changing the total amount of nonlocality.

In a relativistic theory, there is no absolute simultaneity. If the two halves of the experiments are mutually space-like, either one of them can be considered as 'first' and subject to equations (3), while the other one is 'last' and may violate (3). This leads us to speculate that if we want to have a relativistic quantum theory with hidden variables, the values of the latter cannot be Lorentz invariant.

In the foregoing, we have assumed that the orientations of the analysers are set up *independently* and that no information can be conveyed between them by conventional means, as in the experiment proposed by Aspect (1976). This property, however, does not hold for ordinary experiments where the orientations of the analysers are prepared well in advance. In that case, the violation of Bell's inequality by quantum theory has different implications than when the orientations are really independent: the correlation may be construed as due to a 'common cause' proceeding with subluminal velocity. It is noteworthy that quantum theory has *never* been tested in the regime considered by Aspect: usually, experiments are performed in macroscopic space–time regions which are extremely elongated in the time direction. A space-like experiment is certainly very desirable.

Acknowledgment

This research was supported by a grant from the Fund for the encouragement of research at the Technion.

References

Aspect A 1976 *Phys. Rev.* D **14** 1944
Bell J S 1964 *Physics* **1** 195
—— 1966 *Rev. Mod. Phys.* **38** 447
Bohm D and Bub J 1966 *Rev. Mod. Phys.* **38** 453
Clauser J F 1976 *Phys. Rev. Lett.* **36** 1223
Clauser J F and Shimony A 1978 *Rep. Prog. Phys.* **41** 1881
Costa de Beauregard O 1978 *Phys. Lett.* A **67** 171
—— 1979a *Nuovo Cim.* B **51** 267
—— 1979b *Lett. Nuovo. Cim.* **25** 91
Davidon W C 1977 *Nuovo Cim.* B **36** 34
Eberhard P 1977 *Nuovo Cim.* B **38** 75
Einstein A 1949 in *Albert Einstein, Philosopher-Scientist* ed P A Schilpp (Library of Living Philosophers, Evanston) p 85
Freedman S J and Clauser J F 1972 *Phys. Rev. Lett.* **28** 938
Fry E S and Thompson R C 1976 *Phys. Rev. Lett.* **37** 465
Herbert N 1975 *Am. J. Phys.* **43** 315
Peres A 1978 *Am. J. Phys.* **46** 745
Slaby M 1978 *Bull. Am. Phys. Soc.* **23** 586
Stapp H P 1971 *Phys. Rev.* D **3** 1303
—— 1977 *Nuovo Cim.* B **40** 191