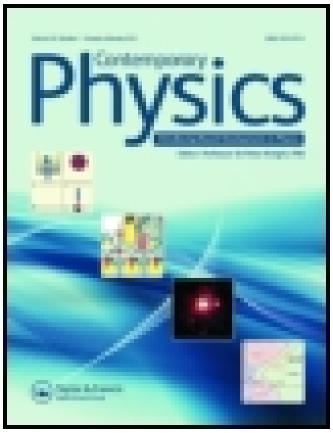

Contemporary Physics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tcph20

Spooky action at a distance in quantum mechanics

Lucien Hardy

Published online: 08 Nov 2010.

To cite this article: Lucien Hardy (1998) Spooky action at a distance in quantum mechanics, Contemporary Physics, 39:6, 419-429, DOI: 10.1080/001075198181757

To link to this article: http://dx.doi.org/10.1080/001075198181757

Contemporary Physics, 1998, volume 39, number 6, pages 419–429

Spooky action at a distance in quantum mechanics

LUCIEN HARDY

In this article we give an introduction to non-locality in quantum mechanics. A simple summary of the basic quantum mechanics required is given. Then we present Bell's proof of quantum non-locality and also we give simple presentations of more recent proofs. Finally we consider the difficulties faced by experimentalists in testing quantum non-locality and review some of the many experiments that have been performed.

1. Introduction

Quantum mechanics is the theory used to make predictions about small systems such as atoms, electrons, photons, etc. In fact, since bigger systems are made out of smaller systems, quantum mechanics also applies to big systems as well. Although the theory is very successful, it is cast in a form which is very different from classical theories (such as Newtonian mechanics and classical electromagnetism) which preceded it. Classical theories point to a rather clear picture of the world. However, the mathematical formulation of quantum mechanics does not point towards a clear picture of the world. There are many different ways of interpreting it. All these interpretations have strange features. In this article we will be interested in one such strange feature, namely an apparent 'spooky' action at a distance, otherwise known as non-locality, which seems to be implied by the predictions of quantum mechanics. Although hinted at by Einstein in 1927 [1], this feature was finally proven by Bell in 1964 [2]. We will go through Bell's proof and also through simpler proofs which have appeared more recently.

2. The story of Albert and Betty

Let us begin with a story. Albert and Betty are (non-identical) twins who earn their living on stage entertaining an audience of scientists every evening. Their stage-act takes the following form. They come on the stage and spend a few minutes conferring secretly together. Next, they are taken to opposite ends of the stage and, by means of screens, members of the audience ensure that they have absolutely no way of communicating with each other. Two members of the audience are nominated as questioners. One stands next to Albert and the other stands next to Betty. Each questioner takes a coin out of their pocket and tosses it. If it comes up heads they ask the 'colour' question: what is your favourite colour? The twin must answer green or red. If the coin comes up tails then the questioner asks the 'food' question: what is your favourite food? The twin must answer carrots or peas. The scientists are careful to ensure that neither twin knows which question is being asked to the other twin. Records of the answers given by the twins each evening are kept. Although the twins often change their mind (so that, for example, on some evenings Albert prefers peas and on others he prefers carrots), the following patterns nevertheless emerge.

(a) On those evenings when both twins are asked the colour question they sometimes (in 9% of cases) both give the answer 'green'.

(b) If one twin is asked the colour question and gives the answer 'green' and the other twin is asked the food question then he/she will always give the answer 'peas'. (Green for one twin implies peas for the other.)

(c) On those evenings when both twins are asked the food question it never happens that they both give the answer 'peas' (one or both of them will answer 'carrots').

The audience is astounded since they realize that this should not be possible. To see why consider how the twins might try to ensure that their answers conform to the above rules. When they come out onto stage they do not know what question they are later going to be asked. It would seem that the best thing they can do is to agree between themselves every evening how they are going to answer each question if it is asked. On some evenings they must both agree to answer green if the colour question is asked in

Author's address: The Clarendon Laboratory, Oxford University, Oxford OX1 3PU, England.

order to satisfy (a). Consider such an evening. We can imagine that while conferring secretly, they each take out a piece of paper. Under the heading 'colour' they each write 'green'. Now Albert notices that since Betty is going to answer green if asked the colour question, he must answer peas if asked the food question to be sure of satisfying (b). Thus he writes 'peas' under the heading 'food'. Likewise, Betty, noticing that Albert has written green, must write 'peas' under her 'food' heading. However they now have a problem. If they are both asked the food question (and they have no way of knowing whether this will happen) they will both answer peas. But this will violate (c). This method of exchanging instruction codes simply cannot work.

The scientists worry that despite their efforts, the twins have some secret way of communicating. If they had been able to sneak in mobile phones for example then they could successfully satisfy the above rules. Albert could simply phone Betty if he has been asked the food question and then she would know not to answer peas if she is also asked the food question. Indeed, the more the scientists discuss this, the more they are convinced that the twins are communicating. To prevent this they obtain a much longer stage. So long that it takes light several minutes to transverse the stage. They reason that since any signals used for communication must travel at a speed less than or equal to that of light, they can prevent the twins from communicating by asking them the questions simultaneously at opposite ends of this rather long stage. This way each twin would have to answer their question before they could possibly find out what the other twin had been asked. The show goes on and after a while it is found that the twins are still able to satisfy the above rules.

Could this happen? If it could it would seem to imply that the twins had found a way of communicating faster than light. Surely then we must conclude that this is prohibited by physics (namely special relativity) and therefore could not happen. But we would be wrong. In fact, by employing quantum mechanics, Albert and Betty could satisfy the above rules. Later I will show exactly how this is possible. However, even though it is possible, Albert and Betty are not actually able to send superluminal signals to one another because there is an uncontrollable randomness in the answers they will give if they use a quantum mechanical state to decide what these should be. This means special relativity is safe (using the words of Shimony it is probably better to say there is a 'peaceful coexistence' between quantum mechanics and special relativity [3]). Before discussing this non-locality I will introduce some basic quantum mechanics.

3. Some basic quantum mechanics

In fact the quantum mechanics required is rather simple. In any physical theory it is necessary to have a way of describing the physical systems that the theory treats. In quantum mechanics a system is described by a vector in a complex linear vector space. In such a space two vectors can each be multiplied by a complex number and added together to give a new vector in the space. Actually, in the examples we will consider, real numbers will be sufficient. We illustrate our remarks by considering the polarization degree of freedom of photons. A photon with vertical polarization will be represented by the vector $|\updownarrow\rangle$. (This notation $|\rangle$ is standard in quantum mechanics for representing vectors and was introduced by Dirac. Readers unfamiliar with this notation should not be put off. We could just as well use a standard vector notation like **a**.) A photon with horizontal polarization will be represented by the vector $|\leftrightarrow\rangle$. To find out whether a photon has vertical or horizontal polarization a calcite crystal can be used. This crystal has an axis which can be orientated along say the vertical direction. Then a vertically polarized photon incident on the crystal will come out along the undeflected path (called the ordinary path) and a horizontally polarized photon will come out along the deflected path (called the extraordinary path). With such a crystal it is possible to distinguish with certainty between photons which are either vertically or horizontally polarized. States which can be distinguished with certainty are said to be orthogonal. (It so happens that orthogonally polarized photons correspond to orthogonal directions in space, e.g. vertical and horizontal. With spin-half particles this is not the case.) A photon with linear polarization along a direction θ to the vertical (see figure 1) is represented by the state

$$|\theta\rangle = \cos(\theta)|\updownarrow\rangle + \sin(\theta)|\leftrightarrow\rangle\,. \tag{1}$$

A photon with orthogonal polarization to this has the state

$$|\theta^{\perp}\rangle = \sin(\theta)|\updownarrow\rangle - \cos(\theta)|\leftrightarrow\rangle\,. \tag{2}$$

By orientating the axis of a calcite crystal at angle θ to the vertical it would be possible to distinguish photons prepared in one or the other of these two states. It is possible to invert equations (1) and (2) to obtain

$$|\updownarrow\rangle = \cos(\theta)|\theta\rangle + \sin(\theta)|\theta^{\perp}\rangle\,, \tag{3}$$

$$|\leftrightarrow\rangle = \sin(\theta)|\theta\rangle - \cos(\theta)|\theta^{\perp}\rangle\,, \tag{4}$$

These equations will be useful later.

Imagine now that we have two photons which have been prepared separately and have not interacted. Photon 1 has been prepared in the state $|\theta\rangle_1$ and photon 2 has been prepared in the state $|\phi\rangle_2$. The combined state of the system is represented by the vector

$$|\theta\rangle_1|\phi\rangle_2\,. \tag{5}$$

Such a state is called a product state since it is possible to represent the state of the combined system of the two particles by a simple product of states of each of the individual particles. However, if the photons have interacted or been prepared in the same process then, in general, the state of the two photons cannot be written as a product. Rather it must be written as a sum of product terms. For example, representing the state of the combined system by the single vector $|\Psi\rangle$, we could have

$$|\Psi\rangle = a|\theta\rangle_1|\phi\rangle_2 + b|\theta'\rangle_1|\phi'\rangle_2 , \tag{6}$$

where a and b are complex numbers. When the state cannot be written as a product it is called an entangled state. It is entangled states that have the almost magical properties that can lead to apparent faster than light influences. When two particles are in an entangled state they appear to continue to talk to each other even after they have finished interacting directly.

To complete this basic introduction to quantum mechanics I want to explain how to calculate probabilities of getting certain outcomes when measurements are made, for example, using calcite crystals. Suppose a photon is prepared with linear polarization ϕ. This means that its state is

$$|\phi\rangle = \cos(\phi)|\updownarrow\rangle + \sin(\phi)|\leftrightarrow\rangle . \tag{7}$$

Now consider letting this photon pass through a calcite crystal orientated at an angle θ (see figure 2). This photon might go into the ordinary channel which corresponds to the state $|\theta\rangle$ or it might go into the extraordinary channel which corresponds to the state $|\theta^{\perp}\rangle$. Using equations (3) and (4) we can write

$$\begin{aligned}|\phi\rangle &= \cos(\phi)[\cos(\theta)|\theta\rangle + \sin(\theta)|\theta^{\perp}\rangle] \\ &\quad + \sin(\phi)[\sin(\theta)|\theta\rangle - \cos(\theta)|\theta^{\perp}\rangle]\end{aligned} \tag{8}$$

rearranging this gives

$$|\phi\rangle = \cos(\theta - \phi)|\theta\rangle + \sin(\theta - \phi)|\theta^{\perp}\rangle . \tag{9}$$

The coefficient in front of the $|\theta\rangle$ term is the probability amplitude associated with the state $|\theta\rangle$. According to the rules of quantum mechanics the probability that the photon will be detected in the ordinary channel is given by the square of the modulus of this probability amplitude, i.e. $\cos^2(\theta - \phi)$. The probability that it will be detected in the extraordinary channel is given by $\sin^2(\theta - \phi)$. Thus, we see the trick is to write down the state using the orthogonal states that correspond to the measurement being performed. The probabilities are given by taking the squares of moduli of the corresponding coefficients. It is important that the probabilities add up to one. This is arranged by normalizing the state so that the length of the vector is equal to 1.

We will now spell out how to calculate probabilities for two photons. Thus, imagine that two photons are prepared in the state

$$|\Psi\rangle = \alpha|\updownarrow\rangle_1|\leftrightarrow\rangle_2 - \beta|\leftrightarrow\rangle_1|\updownarrow\rangle_2 . \tag{10}$$

The state must be normalized to length 1. This means that

$$|\alpha|^2 + |\beta|^2 = 1 . \tag{11}$$

We let the two photons separate to some large distance and then perform measurements of polarization on each one using calcite crystals orientated at angle θ for photon 1 and ϕ for photon 2 (see figure 3). There are two possible outcomes at each end making four possible outcomes in total. To calculate the probabilities of these we first rewrite $|\Psi\rangle$ in terms of $|\theta\rangle_1$, $|\theta^{\perp}\rangle_1$, $|\phi\rangle_2$ and $|\phi^{\perp}\rangle_2$. Using (3) and (4) as before we obtain

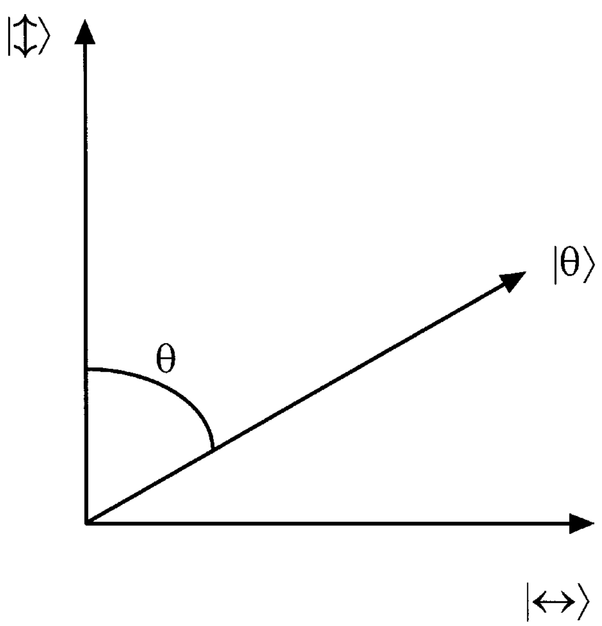

Figure 1. The state vector of a photon polarized at angle θ to the vertical can be resolved into vertical and horizontal polarization vectors.

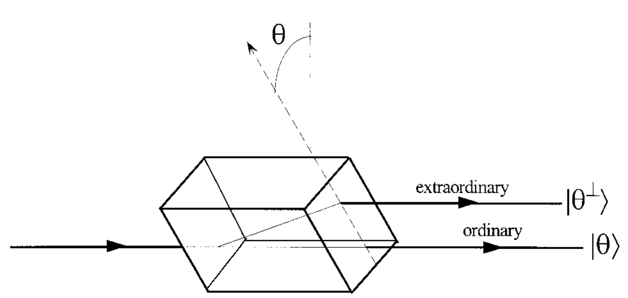

Figure 2. The effect of a calcite crystal orientated with its axis at an angle θ. The ordinary ray corresponds to the state $|\theta\rangle$ while the extraordinary ray corresponds to the state $|\theta^{\perp}\rangle$.

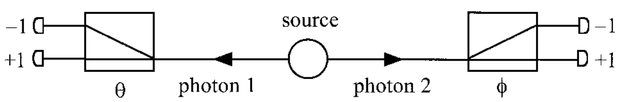

Figure 3. Two entangled photons, 1 and 2, are emitted from the source and impinge on calcite crystals orientated at angles θ and ϕ respectively.

$$|\Psi\rangle = \alpha[\cos(\theta)|\theta\rangle_1 + \sin(\theta)|\theta^{\perp}\rangle_1][\sin(\phi)|\phi\rangle_2 - \cos(\phi)|\phi^{\perp}\rangle_2] - \beta[\sin(\theta)|\theta\rangle_1 - \cos(\theta)|\theta^{\perp}\rangle_1][\cos(\phi)|\phi\rangle_2 + \sin(\phi)|\phi^{\perp}\rangle_2] . \quad (12)$$

Rearranging this we obtain

$$\begin{aligned}|\Psi\rangle = {} & [\alpha\cos(\theta)\sin(\phi) - \beta\sin(\theta)\cos(\phi)]|\theta\rangle_1|\phi\rangle_2 \\ & + [\alpha\cos(\theta)\cos(\phi) - \beta\sin(\theta)\sin(\phi)]|\theta\rangle_1|\phi^{\perp}\rangle_2 \\ & + [\alpha\sin(\theta)\sin(\phi) + \beta\cos(\theta)\cos(\phi)]|\theta^{\perp}\rangle_1|\phi\rangle_2 \\ & + [-\alpha\sin(\theta)\cos(\phi) + \beta\cos(\theta)\sin(\phi)]|\theta^{\perp}\rangle_1|\phi^{\perp}\rangle_2 .\end{aligned} \quad (13)$$

From this equation it is now very easy to write down the probabilities of each of the four possible outcomes. For example, the probability of photon 1 being detected in the ordinary channel (corresponding to $|\theta\rangle_1$) and photon 2 being detected in the extraordinary channel (corresponding to $|\phi^{\perp}\rangle_2$) is given by the square of the modulus of the probability amplitude of the second term:

$$\text{prob}(\theta, \phi^{\perp}) = |\alpha\cos(\theta)\cos(\phi) - \beta\sin(\theta)\sin(\phi)|^2 . \quad (14)$$

Probabilities for the other outcomes are given by taking the square of the modulus of the coefficient in front of the corresponding term. This is all the technical knowledge required to understand the rest of this article.

4. Some early concerns about quantum mechanics

We will take a historical approach to the subject. After the rules of quantum theory had been established in the late 1920s, people began to worry about their physical interpretation. At the 1927 Solvay conference Einstein presented an argument [1] which showed that either quantum mechanics is incomplete or it is non-local. A simplified version of this argument is as follows. A box has a particle in it. The box is divided by means of partitions and split into two separate boxes (A and B). These two boxes are then taken to two remote places (see figure 4). The state of the system at this stage can be written

$$|\psi\rangle = \frac{1}{2^{1/2}}(|A\rangle + |B\rangle) \quad (15)$$

where $|A\rangle$ ($|B\rangle$) denotes that the particle is in box A (B). This state vector illustrates the superposition principle.

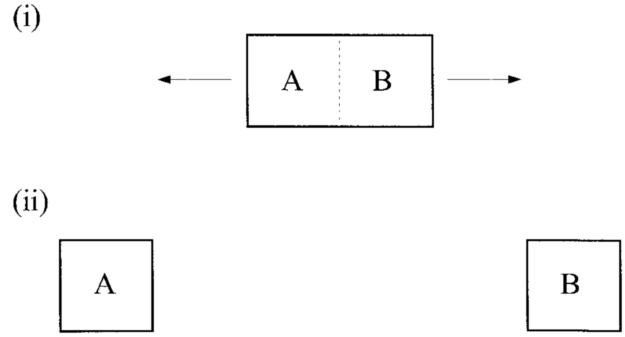

Figure 4. (i) A particle is placed into a box which is then partitioned into two boxes. (ii) The two boxes are separated to a great distance. Is the particle always in one of the two boxes or is it the case that there is no fact of the matter as to which box the particle is in until one of them is opened?

Generally, when we do not know which of two possibilities has actually happened, the state is written as a sum of the corresponding vectors. Now box A is opened. There is a 50% chance that the particle will be found there. If the particle is found there then we know that it will not be found in the other box. There are two ways we can think about this thought experiment: (i) the particle is actually always in one of the two boxes (box A say). When box A is opened it is simply discovered there. If we take this approach then we must believe that there is some additional information about the system, namely which box the particle is in for a particular run of the experiment, that is not included in the state description given above in (15). That is to say we must accept that the quantum mechanical description of the state is incomplete. Alternatively we can assert that (15) does represent a complete description of the state of affairs. Then we tell the second story: (ii) there is no fact of the matter as to which box the particle is in before one of them is opened. The moment box A is opened it becomes the case that the particle is in box A (assuming it is found to be there). If we believe this then there must be a non-local influence from box A to box B carrying the information that since the particle has been found in box A it cannot also be found in box B. Without such a non-local influence it could happen that when box B is opened the particle is also found there and one particle should not be found to be in two places at the same time. Both stories are problematic. Either we must accept that quantum mechanics is incomplete, or we must accept that there are non-local influences.

This dilemma was presented again later in a different form in a very famous paper by Einstein, Podolsky, and Rosen (EPR) in 1935 [4]. This paper is very easy to read and I strongly recommend it to the reader. The example considered by EPR consisted of two particles in an entangled state involving their positions and momenta but, following Bohm [5], we will illustrate their argument by considering two photons in the following polarization entangled state (actually Bohm considered two spin-half particles but two photons are equivalent for our purposes):

$$|\Psi\rangle = \frac{1}{2^{1/2}}(|\updownarrow\rangle_1|\leftrightarrow\rangle_2 - |\leftrightarrow\rangle_1|\updownarrow\rangle_2) \, . \tag{16}$$

You can easily verify that this state is rotationally invariant in form in the sense that if it is written in terms of the vectors $|\theta\rangle_{1,2}$ and $|\theta^{\perp}\rangle_{1,2}$ then it has the form

$$|\Psi\rangle = \frac{1}{2^{1/2}}(|\theta\rangle_1|\theta^{\perp}\rangle_2 - |\theta^{\perp}\rangle_1|\theta\rangle_2) \, . \tag{17}$$

This means that if a measurement of polarization is made on both photons along the same direction θ and one photon is detected in the ordinary path, the other photon will certainly be detected in the extraordinary path and vice versa (since the terms $|\theta\rangle_1|\theta\rangle_2$ and $|\theta^{\perp}\rangle_1|\theta^{\perp}\rangle_2$ have zero amplitude). This is like the particle in one of two boxes. We have similar choice. We can imagine either that the quantum description (17) is complete and hence when a measurement is made on photon 1 there is a non-local influence letting photon 2 know the outcome. Or we can imagine that before the two photons separate they agree between themselves that should the polarization be measured along the direction θ one photon should come out in the θ channel and the other should come out in the $\theta^{\perp}$ channel so we preserve locality but now the quantum state is not taken to be a complete description of what is going on.

5. Enter John S. Bell

The second of these two options amounts to the photons exchanging an instruction code before separating specifying how they will behave when a given measurement is made. Such instruction codes go beyond the normal quantum state description of a system and are usually called hidden variables. We will see that this attempt to rescue locality by introducing hidden variables fails when any attempt is made to put it into action (indeed, Albert and Betty's trick depends on its failure). An early hint that it would fail was provided by the de Broglie Bohm interpretation of quantum mechanics. At the 1927 Solvay conference de Broglie [6] had put forward an interpretation of quantum mechanics which explicitly involved the introduction of hidden variables. This interpretation was rediscovered and further developed by Bohm in 1952 [7]. One feature of this interpretation is that it is explicitly non-local. Indeed, it was this feature that lead Einstein to reject the interpretation. However, John Bell asked if this non-locality was a necessary feature of any attempt to supplement quantum mechanics with hidden variables. He was subsequently (in 1964) able to prove that it is [2].

Bell considered the state in equation (16) (actually, like Bohm, Bell also considered two spin-half particles but, again, two photons are equivalent for our purposes) which is the same as equation (10) if we put $\alpha = \beta = 2^{-1/2}$. Further, he defined the correlation function

$$E(\theta, \phi) = \text{Average}\,(A(\theta)B(\phi)) \, , \tag{18}$$

where $A(\theta) = +1$ if photon one comes out in the ordinary channel (corresponding to $|\theta\rangle_1$) and $A(\theta) = -1$ if it comes out in the extraordinary channel (corresponding to $|\theta^{\perp}\rangle_1$) and where $B(\phi)$ is defined similarly for photon 2 (as in figure 3). The average is taken over a large number of runs of the experiment. The probabilities $\text{prob}\,(\theta, \phi)$ and $\text{prob}\,(\theta^{\perp}, \phi^{\perp})$ contribute to $A(\theta)B(\theta) = +1$ and the probabilities $\text{prob}\,(\theta, \phi^{\perp})$ and $\text{prob}\,(\theta^{\perp}, \phi)$ contribute to $A(\theta)B(\theta) = -1$. Hence, the correlation function defined above is equal to

$$\begin{aligned} E(\theta, \phi) &= \text{prob}\,(\theta, \phi) + \text{prob}\,(\theta^{\perp}, \phi^{\perp}) \\ &\quad - \text{prob}\,(\theta, \phi^{\perp}) - \text{prob}\,(\theta^{\perp}, \phi) \, . \end{aligned} \tag{19}$$

The quantum predictions for these probabilities can be read off from equation (13) (putting $\alpha = \beta$). After a little trigonometric manipulation this gives

$$E(\theta, \phi) = -\cos[2(\theta - \phi)] \, . \tag{20}$$

Now consider thinking about this experiment from the point of view of locality. Imagine that before the two photons separate they exchange an instruction code as we discussed above. This instruction code can be represented by λ and are called hidden variables. These hidden variables can be anything we want. They can, and in general will, be different from one run of the experiment to the next. We imagine that, over a large number of runs of the experiment, they are distributed according to some probability density function $\rho(\lambda)$. Since this is a probability density, it has the property

$$\int \rho(\lambda)\, \mathrm{d}\lambda = 1 \, , \tag{21}$$

where the integration is over the space of λ. The values of $A(\theta)$ and $B(\phi)$ for a given run of the experiment will be determined by λ so we can write the result functions $A(\theta, \lambda)$ and $B(\phi, \lambda)$. Note in particular that these result functions do not depend on the setting at the other end of the experiment. This is equivalent to assuming that Albert and Betty are not allowed to communicate and forms the basis of the locality assumption. Further note that the distribution $\rho(\lambda)$ does not depend on $\theta_{1,2}$. This also forms part of the locality assumption and is equivalent to the fact that Albert and Betty do not know what questions they are going to be asked. It follows that in a local hidden variable theory the correlation function can be written as

$$E(\theta, \phi) = \int A(\theta, \lambda) B(\phi, \lambda) \rho(\lambda)\, d\lambda \, . \tag{22}$$

We will see that this form for the correlation function is inconsistent with the quantum result (20).

First we derive a simple mathematical result. Assume $x_1, x_1', x_2, x_2' = \pm 1$. Then

$$s = x_1 x_2 + x_1 x_2' + x_1' x_2 - x_1' x_2' = \pm 2 \, . \tag{23}$$

This is most easily seen by writing

$$s = x_1(x_2 + x_2') + x_1'(x_2 - x_2') \, . \tag{24}$$

One of the two expressions in brackets must be zero and it follows that $s = \pm 2$.

Now we put $x_1 = A(\theta, \lambda)$, $x_1' = A(\theta', \lambda)$, $x_2 = B(\phi, \lambda)$ and $x_2' = B(\phi', \lambda)$ into (23), multiply by $\rho(\lambda)$ and then integrate over λ. This gives

$$-2 \le E(\theta, \phi) + E(\theta, \phi') + E(\theta', \phi) - E(\theta', \phi') \le +2 \, . \tag{25}$$

These are called Bell inequalities (actually this version of Bell inequalities were derived by Clauser *et al.* [8]). They express a constraint that all local hidden variable theories must obey. Now use $\theta = 0^\circ$, $\theta' = 45^\circ$, $\phi = 22.5^\circ$, and $\phi' = -22.5^\circ$ as shown in figure 5. If these values are inserted into (25) using the quantum mechanical formula (20) then we obtain

$$S = E(\theta, \phi) + E(\theta, \phi') + E(\theta', \phi) - E(\theta', \phi') = 2(2^{1/2}) \, . \tag{26}$$

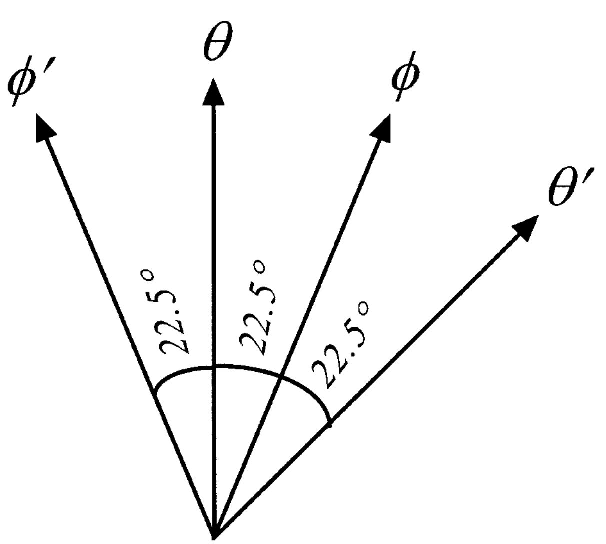

Figure 5. The settings for the polarization measurements used to show a violation of Bell's inequalities.

This violates the Bell inequalities which means that the quantum predictions cannot be reproduced by a local hidden variable model. This result is quite astonishing. What it means is that there must be a non-local influence passing between the two particles. Indeed, this is so astonishing that many people did not believe that quantum mechanics could be right. It was necessary to do experiments. These will be described later.

6. Proofs without inequalities

A number of more direct proofs of quantum non-locality without appeal to inequalities were discovered in the 1980s and 1990s. The first of these was due to Heywood and Redhead [9] which pertained to two spin one particles. Their proof was simplified by Stairs [10] and further simplified by Brown and Svetlichny [11] although it is still rather complicated. A much simpler proof which we will present below was due to Greenberger, Horne and Zeilinger (GHZ) [12] and employed three or more particles. In 1991, motivated by the proof of GHZ I came up with a rather simple proof [13] for two photons (or spin half particles). It is this that forms the basis of the trick of Albert and Betty and it will also be presented below. More recently, Penrose [14] has come up with another example which uses two spin one particle like that of Heywood and Redhead.

7. The Greenberger–Horne–Zeilinger proof

We will now present the Greenberger–Horne–Zeilinger (GHZ) [12] proof of non-locality without inequalities. Their proof requires more than two particles. We will give a version of the argument employing three photons [15]. We will use the notation

$$|+\rangle \equiv |\updownarrow\rangle \qquad |-\rangle \equiv |\leftrightarrow\rangle \tag{27}$$

for vertically and horizontally polarized states and

$$|+'\rangle \equiv |\theta = +45^\circ\rangle = \frac{1}{2^{1/2}}(|+\rangle + |-\rangle) \, , \tag{28}$$

$$|-'\rangle \equiv |\theta = -45^\circ\rangle = \frac{1}{2^{1/2}}(|+\rangle - |-\rangle) \, , \tag{29}$$

for states polarized at $\pm 45^\circ$ to the vertical (we have used equations (1) and (2)).

Before going on to the main result we mention that the following equations follow from (28) and (29)

$$\frac{1}{2^{1/2}}(|+\rangle_2|+\rangle_3 - |-\rangle_2|-\rangle_3) = \frac{1}{2^{1/2}}(|+'\rangle_2|-'\rangle_3 + |-'\rangle_2|+'\rangle_3) \, , \tag{30}$$

$$\frac{1}{2^{1/2}}(|+\rangle_2|-\rangle_3 + |-\rangle_2|+\rangle_3) = \frac{1}{2^{1/2}}(|+'\rangle_2|+'\rangle_3 - |-'\rangle_2|-'\rangle_3) \, . \tag{31}$$

These results will be used later (note, we have written them down for photons 2 and 3 since this example will be used later but the result holds equally for any other pair of photons).

The three photons are prepared in a source in the state

$$|\Psi\rangle = \frac{1}{4^{1/2}}(|+\rangle_1|+\rangle_2|+\rangle_3 - |+\rangle_1|-\rangle_2|-\rangle_3 - |-\rangle_1|+\rangle_2|-\rangle_3 - |-\rangle_1|-\rangle_2|+\rangle_3) . \quad (32)$$

The photons are allowed to propagate to three separate places where measurements of polarization are made on each photon as shown in figure 6. At each of these three locations a random choice is made to measure either the polarization along the vertical and horizontal directions or along the $\pm 45^\circ$ directions. Let the polarization of photon 1 be measured along a direction at angle θ to the vertical. We set $A(\theta) = +1$ if it is detected in the θ channel and $A(\theta) = -1$ if it is detected in the $\theta^\perp$ channel. The quantities $B(\phi)$ and $C(\chi)$ are defined in a similar way for photons 2 and 3 for measurements of polarization at angles ϕ and χ to the vertical respectively. Consider the case in which $(A(0^\circ), B(0^\circ), C(0^\circ))$ are measured. It follows from (32) that the only possibilities for the results of these three measurements are $(+1, +1, +1)$, $(+1, -1, -1)$, $(-1, +1, -1)$, $(-1, -1, +1)$. The product of the three outcomes is + 1 in each case. Hence we can say that

$$A(0^\circ)B(0^\circ)C(0^\circ) = +1 . \quad (33)$$

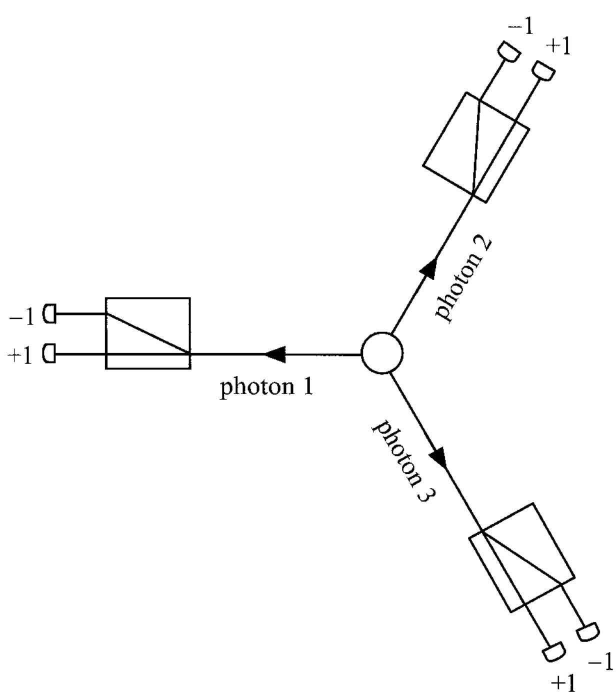

Figure 6 The apparatus used in the Greenberger–Horne–Zeilinger argument. Three photons prepared in a particular state are emitted from the source. Measurements of polarization are made on each photon with the corresponding calcite crystal orientated either at 0° or 45°.

Now we will consider the case in which $A(0^\circ)$ is measured on photon 1 and $B(45^\circ)$ and $C(45^\circ)$ are measured on photons 2 and 3 respectively. First notice that (32) can be written

$$|\Psi\rangle = \frac{1}{2^{1/2}}|+\rangle_1[\frac{1}{2^{1/2}}(|+\rangle_2|+\rangle_3 - |-\rangle_2|-\rangle_3)] - \frac{1}{2^{1/2}}|-\rangle_1[\frac{1}{2^{1/2}}(|+\rangle_2|-\rangle_3 + |-\rangle_2|+\rangle_3)] . \quad (34)$$

Using equations (30) and (31) we can write this as

$$|\Psi\rangle = \frac{1}{2^{1/2}}|+\rangle_1[\frac{1}{2^{1/2}}(|+'\rangle_2|-'\rangle_3 + |-'\rangle_2|+'\rangle_3)] - \frac{1}{2^{1/2}}|-\rangle_1[\frac{1}{2^{1/2}}(|+'\rangle_2|+'\rangle_3 - |-'\rangle_2|-'\rangle_3)] . \quad (35)$$

This can be multiplied out to give

$$|\Psi\rangle = \frac{1}{4^{1/2}}(|+\rangle_1|+'\rangle_2|-'\rangle_3 + |+\rangle_1|-'\rangle_2|+'\rangle_3 - |-\rangle_1|+'\rangle_2|+'\rangle_3 + |-\rangle_1|-'\rangle_2|-'\rangle_3) . \quad (36)$$

Hence when the measurements $(A(0^\circ), B(45^\circ), C(45^\circ))$ are made the possible results are $(+1, +1, -1)$, $(+1, -1, +1)$, $(-1, +1, +1)$, $(-1, -1, -1)$. In each case the product of the three outcomes is equal to -1. That is

$$A(0^\circ)B(45^\circ)C(45^\circ) = -1 . \quad (37)$$

By symmetry two further properties also hold

$$A(45^\circ)B(0^\circ)C(45^\circ) = -1 , \quad (38)$$

$$A(45^\circ)B(45^\circ)C(0^\circ) = -1 . \quad (39)$$

We can now think about how to reproduce these properties in a local hidden variable model. As before we can introduce the hidden variables λ which represent the instruction code that is shared by the photons before leaving the source. Each result will be determined by a result function like $A(0^\circ, \lambda)$. It follows from equations (36), (37), (38) and (39) that

$$A(0^\circ, \lambda)B(0^\circ, \lambda)C(0^\circ, \lambda) = +1 , \quad (40)$$

$$A(0^\circ, \lambda)B(45^\circ, \lambda)C(45^\circ, \lambda) = -1 , \quad (41)$$

$$A(45^\circ, \lambda)B(0^\circ, \lambda)C(45^\circ, \lambda) = -1 , \quad (42)$$

$$A(45^\circ, \lambda)B(45^\circ, \lambda)C(0^\circ, \lambda) = -1 . \quad (43)$$

Now consider multiplying these four equations. The product of the right hand sides is -1. However, on the left hand side each quantity appears twice. Since each quantity is equal to ± 1 the product of the left hand sides is

$+1$. This gives a direct contradiction between the predictions of quantum theory and local hidden variable models. The above four equations simply cannot all be true. It is worth noting that dropping the locality assumption removes the contradiction. For example consider the quantity $A(0^\circ, \lambda)$ in equations (40) and (41). If this quantity is allowed to depend on what is being measured on the other two photons then it could take different values in (40) and (41). This would remove the contradiction.

8. How Albert and Betty's trick works

So far we have not explained how Albert and Betty did the trick discussed at the beginning of this article. In 1992, motivated by the GHZ argument and employing some arguments from a paper by Elitzur and Vaidman [16], I showed how it is possible to realize a simple two particle argument for non-locality without inequalities [13]. The correlations are exactly those needed in Albert and Betty's trick. We will now describe exactly what Albert and Betty need to do. They require two photons 1 and 2. Each photon lives in a two-dimensional vector space spanned by the orthogonal vectors $|\updownarrow\rangle$ and $|\leftrightarrow\rangle$. This vector space can also be spanned by two differently orientated orthogonal polarization vectors which we will call $|r\rangle$ and $|g\rangle$ corresponding to 'red' and 'green'. Actually, it does not matter which way they are orientated, so long as they are orthogonal. Now we introduce a further two orthogonal vectors $|p\rangle$ and $|c\rangle$ (corresponding to 'peas' and 'carrots') which are orientated at some angle to $|r\rangle$ and $|g\rangle$ (see figure 7) such that

$$|r\rangle = a|p\rangle + b|c\rangle \,, \tag{44}$$

$$|g\rangle = b|p\rangle - a|c\rangle \,, \tag{45}$$

where we are taking the coefficients a and b to be real numbers. The vectors are taken to be normalized so that $a^2 + b^2 = 1$. These two equations can be inverted to give

$$|p\rangle = a|r\rangle + b|g\rangle \,, \tag{46}$$

$$|c\rangle = b|r\rangle - a|g\rangle \,. \tag{47}$$

When they meet at the centre of the stage Albert and Betty put their two photons in the entangled state

$$|\psi\rangle = N(|r\rangle_1|r\rangle_2 - a^2|p\rangle_1|p\rangle_2) \,, \tag{48}$$

where N is a normalization constant. They then each carry their photon to opposite ends of the stage. When they get there they will be asked a question. If asked the colour question they make a measurement of polarization along the $|r\rangle$, $|g\rangle$ basis. If the outcome is in the $|r\rangle$ channel they answer red and if the outcome is in the $|g\rangle$ channel they answer green. If asked the food question they make a measurement of polarization along the $|p\rangle$, $|c\rangle$ basis. If the outcome is in the $|p\rangle$ channel they answer peas and if the outcome is in the $|c\rangle$ channel they answer carrots. We will now see that their answers will be in accordance with the patterns specified earlier. Consider the case where both Albert and Betty are asked the colour question. They both make measurements on the r, g bases. We can insert equation (46) into (48) to obtain

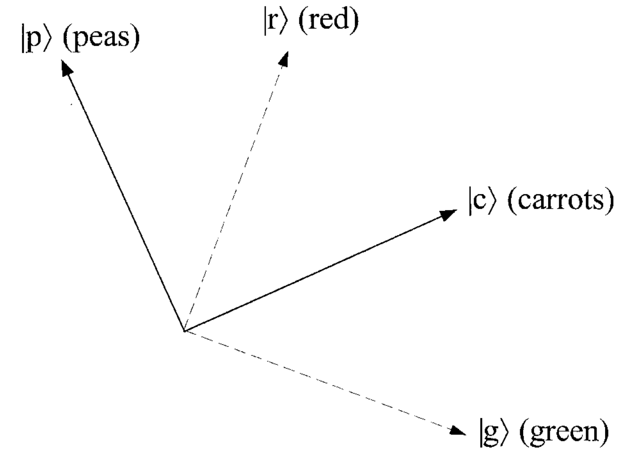

Figure 7. The relationship between the $|p\rangle$ and $|c\rangle$ polarization vectors and the $|r\rangle$ and $|g\rangle$ polarization vectors.

$$|\psi\rangle = N(|r\rangle_1|r\rangle_2 - a^2(a|r\rangle_1 + b|g\rangle_1)(a|r\rangle_2 + b|g\rangle_2)) \,. \tag{49}$$

We can see that the coefficient in front of the $|g\rangle_1|g\rangle_2$ term is Na^2b^2. This means that the probability of both Albert and Betty answering green is non-zero (it is not difficult to show that the maximum value of this probability is about 9%). This satisfies (a) that on those evenings when both twins are asked the colour question they sometimes both give the answer green.

Now consider the case where Albert is asked the colour question and Betty is asked the food question. We can substitute (46) into (48) giving

$$|\psi\rangle = N(|r\rangle_1|r\rangle_2 - a^2(a|r\rangle_1 + b|g\rangle_1)|p\rangle_2) \,. \tag{50}$$

We see that the only term containing $|g\rangle_1$ is the $|g\rangle_1|p\rangle_2$. There is no $|g\rangle_1|c\rangle_2$ term. This means that the probability of Albert answering green and Betty answering carrots is zero. Hence, if Albert answers green then Betty will answer peas. By symmetry it follows that if Betty is asked the colour question and answers green and Bob is asked the food question then he will answer peas. This means that property (b) is satisfied.

Finally, consider those evenings on which both twins are asked the food question. Substitute (44) into (48):

$$|\psi\rangle = N((a|p\rangle_1 + b|c\rangle_1)(a|p\rangle_2 + b|c\rangle_2) - a^2|p\rangle_1|p\rangle_2) \,. \tag{51}$$

Notice that the $|p\rangle_1|p\rangle_2$ term cancels. That is there is zero probability of both twins answering peas on the same evening. This satisfies property (c) and so we have demonstrated that the twins are able to do the trick described earlier.

9. Real experiments

One response to Bell's theorem is that quantum mechanics is wrong—at least in those situations where it predicts non-locality. This is an empirical matter and indeed after Bell's paper was published it was very quickly realized that experiments were necessary. To do this the quantity S defined in equation (26) must be measured. If S were found to have magnitude greater than 2 this would demonstrate that nature cannot be described by a local hidden variable model. However, there is a problem which was first appreciated by Pearle [17]. The detectors used in the experiment will have some efficiency η. In the derivation of the quantum mechanical result (20) we assumed that the detectors are ideal so that $\eta = 1$. However, this is unrealistic in a real experiment. If we take into account the efficiency of the detectors then we obtain

$$E(\theta, \phi) = -\eta^2 \cos[2(\theta - \phi)] \tag{52}$$

since each of the probabilities in equation (19) is multiplied by η^2. This gives $S = 2(2^{1/2})\eta^2$. The Bell inequalities are violated when $S > 2$ which requires

$$\eta > 2^{-1/4} \approx 84\%. \tag{53}$$

This efficiency must take into account not only the detector inefficiency but also other inefficiencies inherent in the experiment such as transmission inefficiency through the polarizer and the collection inefficiencies of the optical elements (not all the photons emitted from the source will enter the optical elements and so cannot arrive at the detectors). By considering different quantum states Eberhard [18] has shown that it is possible to reduce the efficiency required to 67%. Nevertheless, 67% is still a very high efficiency to achieve and to date no experiments to test the Bell inequalities have been performed with such high efficiency.

We can also think about this in the context of the Albert and Betty trick discussed above. In their case, having efficiency less than 100% is equivalent to allowing them to refuse to answer the questions put to them on some occasions. They could refuse to answer questions when they can deduce that answering them could lead to a contradiction of one of the properties (a)–(c). In this way it might be possible to satisfy the properties (a)–(c) in a local way in that proportion of cases where they both answer the questions. For this to be successful they would have to refuse to answer for at least a certain proportion of the questions.

To get round this problem an untestable supplementary assumption was introduced (for a review see [19]). The basic idea of this assumption is that the detectors sample the ensemble in a fair way so that those events in which both photons are detected are representative of the whole ensemble. This is like saying Albert and Betty decide randomly whether or not to answer questions by tossing an appropriately biased coin just after they have been asked the question, such that this cannot form the basis of any strategy. With this assumption we can redefine the correlation function E to be the average value of $A(\theta)B(\phi)$ over those events in which both photons are detected. This is equivalent to renormalizing the original definition so that now we have

$$E(\theta, \phi) = \frac{\text{Average}\,(A(\theta)B(\phi))}{\text{Average}\,(N_1 N_2)}\,, \tag{54}$$

where N_1 is the total number of photons (1 or 0) detected at end 1 and N_2 is defined similarly. Note that

$$\begin{aligned}\text{Average}\,(N_1N_2) = {} & \text{prob}\,(\theta, \phi) + \text{prob}\,(\theta^{\perp}, \phi^{\perp}) \\ & + \text{prob}\,(\theta, \phi^{\perp}) + \text{prob}\,(\theta^{\perp}, \phi)\,.\end{aligned} \tag{55}$$

This means that when inefficiencies are considered, the factor η^2 will enter into both the numerator and the denominator of the RHS of (54) so that it cancels. Hence we recover the prediction (20). Thus, with the fair sampling assumption the Bell inequalities will be violated by the predictions of quantum mechanics regardless of how low the efficiency is. All the experiments performed to date make use of the fair sampling assumption.

Another consideration pointed out by Bell in his original paper is that the choice of which measurement to perform should be made while the photons are in flight to prevent any signals from one end to the other carrying the information of which measurement is to be performed. Most experiments that have been performed are one to two metres long (the length of a typical optical table). To switch between two settings of a polarizer in the time it takes light to travel one metre is very difficult. Nevertheless, one longer experiment has accomplished this though not in an entirely successful way.

The first experiment was performed in 1972 by Freedman and Clauser [20]. They employed an atomic cascade. Atoms were excited in such a way that when they decayed two photons would be produced that were entangled with respect to their polarizations. The experiment violated the Bell inequalities by 6 standard deviations. A number of further experiments employing atomic cascades and other methods were performed in the 1970s and most of them violated the Bell inequalities as expected [21]. Two experiments did not violate the inequalities [22]. Nevertheless, the experiments were overwhelmingly in favour of quantum mechanics. For a review of experiments up to 1978 see [19]. Further experiments by Aspect and co-workers [23] in Orsay in the early 1980s also employing atomic cascades provided further evidence in favour of quantum mechanics. The last of their experiments was 12 m long and employed an active switching device so that the choice of measurement was decided while the photon was in flight. Unfortunately, as was subsequently pointed out by Zeilinger [24], the time period of the switching was chosen such that the switch was back to its original value by the time the photon arrived at the polarizers.

A new generation of experiments started with the discovery of a phenomenon known as parametric down conversion (PMDC). In PMDC a special type of crystal with certain nonlinear optical properties is illuminated by a laser. A small proportion of the photons in the laser beam are 'down converted' into pairs of photons. These pairs of photons can be used to prepare an entangled state. The advantage of this source of entangled photons over atomic cascades are considerable. First, the directions in which the two photons go is well correlated. Secondly, the experiment is considerably easier to set up. Many experiments to test Bell's inequalities employing PMDC have been performed and have shown, as expected, a violation (for a very incomplete list of references see [25]).

To date, no experiment has been performed to test the GHZ correlation because of the difficulty of preparing three photon states although experimental techniques that could lead to an experiment are now being mastered [26] and an experiment is likely to be performed soon. A number of experiments [27] have been performed to test the predictions of quantum mechanics used by Albert and Betty. These experiments agree very well with quantum mechanics.

A very long experiment over 10 km was recently performed by Tittel *et al.* [28]. This employed optical fibres laid down by a telephone company. The source was in Geneva, and the measurements were made in two small villages, Bellevue and Bernex, nearby. Again the Bell inequalities were violated.

The interpretation of all these experiments, that they demonstrate the impossibility of local hidden variable models, depends on believing the fair sampling assumption. Many people question this fair sampling assumption. Indeed, people have shown how it is possible to construct local hidden variable models that can reproduce all the results of performed experiments but which do violate the fair sampling assumption. Presently a number of experimental groups around the world are working towards an experiment with sufficiently high efficiency that the Bell inequalities are violated directly without appeal to this assumption. Given the success of quantum mechanics generally it seems overridingly likely that these experiments will agree with quantum mechanics and violate the Bell inequalities. Nevertheless, opinions formed on such a basis are no substitute for real experimental data. I, for one, am looking forward to the day when this matter is finally settled experimentally.

Acknowledgements

I would like to thank the Royal Society for funding.

References

[1] *See* Jammer, M., 1974, *The Philosophy of Quantum Mechanics* (New York: Wiley), 115–117.

[2] Bell, J. S., 1964, *Physics,* **1,** 195.

[3] Shimony, A., 1984, Controllable and uncontrollable non-locality. *Quantum Mechanics in the Light of New Technology* (Tokyo: The Physical Society of Japan), pp. 225.

[4] Einstein, A., Podolsky, B., and Rosen, N., 1935, *Phys. Rev.,* **47,** 777.

[5] Bohm, D., 1951, *Quantum Theory* (Englewood Cliffs, NJ: Prentice-Hall).

[6] de Broglie, L., 1927, *C. R. Séances Acad. Sci.,* **185,** 380–382; 1928, *Electron. Photon.,* 105–132.

[7] Bohm, D., 1952, *Phys. Rev.,* **85,** 166–179.

[8] Clauser, J. F., Horne, M. A., Shimony, A., and Holt, R. A., 1969, *Phys. Rev. Lett.,* **23,** 880.

[9] Heywood, P., and Redhead, M. L. G., 1983, *Found. Phys.,* **13,** 481.

[10] Stairs, A., 1983, *Philos. Sci.,* **50,** 587.

[11] Brown, H. R., and Svetlichny, G., 1990, *Found. Phys.,* **20,** 1379.

[12] Greenberger, D. M., Horne, M. A., and Zeilinger, A., 1989, *Bell's Theorem, Quantum Theory, and Conceptions of the Universe,* edited by M. Kafatos (Dordrecht: Kluwer); Greenberger, D. M., Horne, M., Shimony, A., and Zeilinger, A., 1990, *Am. J. Phys.,* **58,** 1131.

[13] Hardy, L., 1992, *Phys. Rev. Lett.,* **68,** 2981.

[14] Penrose, R., 1995, *Shadows of the Mind* (UK: Vintage Science).

[15] Mermin, N. D., 1990, *Am. J. Phys.,* **58,** 731.

[16] Elitzur, A., and Vaidman, L., 1993, *Found. Phys.,* **23,** 287.

[17] Pearle, P., 1970, *Phys. Rev.* D, **2,** 1418.

[18] Eberhard, P., 1993, *Phys. Rev.* A, **47,** R747.

[19] Clauser, J. F., and Shimony, A., 1978, *Rep. Prog. Phys.,* **41,** 1881.

[20] Freedman, S. J., and Clauser, J. F., 1972, *Phys. Rev. Lett.,* **28,** 938.

[21] Clauser, J. F., 1976, *Phys. Rev. Lett.,* **36,** 1223; Kasday, L. R., Ullman, J. D., and Wu, C. S., 1975, *Bull. Nuovo Cim.,* **25,** B633; Fry, E. S., and Thompson, R. C., 1976, *Phys. Rev. Lett.,* **37,** 465; Wilson, A. R., Lowe, J., and Butt, D. K., 1976, *J. Phys.* G: *Nucl. Phys.,* **2,** 613; Bruno, M., d'Agostino, M., and Maroni, C., 1977, *Nuovo Cim.,* **40,** B142.

[22] Holt, R. A., and Pipkin, F. M., 1973, preprint of Harvard University; Faraci, G., Gutkowski, S., Notarrigo, S., and Pennisi, A. R., 1974, *Lett. Nuovo Cim.,* **9,** 607.

[23] Aspect, A., Grangier, P., and Roger, G., 1981, *Phys. Rev. Lett.,* **47,** 460; Aspect, A., Grangier, P., and Roger, G., 1982, *Phys. Rev. Lett.,* **48,** 91; Aspect, A., Dalibard, J., and Roger, G., 1982, *Phys. Rev. Lett.,* **49,** 1804.

[24] Zeilinger, A., 1986, *Phys. Lett.* A, **118,** 1.

[25] Alley, C. O., and Shih, Y. H., 1987, *Proceedings of the 2nd International Symposium on the Foundations of Quantum Theory in the Light of New Technology,* edited by M Namiki *et al.* (Tokyo: Physical Society of Japan); Franson, J. D., 1989, *Phys. Rev. Lett.,* **67,** 290; Rarity, J. G., and Tapster, P. R., 1990, *Phys. Rev. Lett.,* **64,** 2495; Kwiat, P. G., Mattle, K., Weinfurter, H., Zeilinger, A., Sergienko, A. V., and Shih, Y., 1995, *Phys. Rev. Lett.,* **75** 4337.

[26] The technology demonstrated in Bouwmeester, D., Pan, J.-W., Mattle, K., Eibl, M., Weinfurter, H., and Zeilinger, A., 1997, *Nature,* **390,** 575 could be used to implement the proposal in Zeilinger, A., Horne, M. A., Weinfurter, H., and Żukowski, M., 1997, *Phys. Rev. Lett.,* **78,** 3031.

[27] Torgerson, J. R., Branning, D., Monken, C.H., and Mandel, L., 1995, *Phys. Lett.* A, **204,** 323; Di Giuseppe, G., De Martini, F., and Boschi, D., 1997, *Phys. Rev.* A, **56**; Boschi, D., Branca, S., De Martini, F., and Hardy, L., 1997, *Phys. Rev. Lett.,* **79,** 2755.

[28] Tittel, W., Brendel, J., Gisin, B., Herzog, T., Zbinden, H., and Gisin, N., 1998, *Phys. Rev.* A, **57,** 3229.

Lucien Hardy took his first degree at Imperial College (1988) and his PhD at Durham University (1992). Since then he has held research and lecturing positions in Ireland (Maynooth), Austria (Innsbruck), Durham and Rome. He now holds a Royal Society University Research Fellowship and is working at the Clarendon Laboratory in Oxford. His research interests include the foundations of quantum theory and quantum information theory.