Nonlocality, violation of lorentz in variance, and wave-particle duality in quantum theory

Hardy, Lucien

Nonlocality, Violation of Lorentz Invariance, and Wave-Particle Duality in Quantum Theory

by

Lucien Hardy

A Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Department of Mathematical Sciences

The University of Durham

1992

Declaration

This thesis is the result of my own work. The work of other authors is referred to in the thesis at the appropriate point in the text.

Abstract

This thesis addresses some of the consequences of giving quantum mechanics a realist interpretation. We consider three main topics: wave-particle duality, locality, and Lorentz invariance.

First we show that classical particles alone or classical waves alone cannot explain all single particle quantum phenomena. Then we consider the possibility that a quantum particle is composed of a particle and a wave, both being taken to exist objectively. We are able to demonstrate the reality of empty waves (that is, waves without a particle) if we make three realist motivated assumptions.

The main part of this thesis concerns locality. In 1964 Bell demonstrated that a local realist interpretation of quantum mechanics is not possible by deriving a set of inequalities that apply to two particle systems. More recently Greenberger, Horne, and Zeilinger have demonstrated this for systems with more than two particles without the need for inequalities. We present a new way to derive Bell inequalities for two particles and show how this can be extended to systems with more than two particles. A number of proposals for experiments to test local realism are put forward. In particular, we show how it is possible to demonstrate the nonlocality of a single photon. A new demonstration of Bell's theorem is presented for two particles but without inequalities. A realizable quantum optical version is proposed and inequalities are proposed which would be required in a non-ideal experiment.

Finally, the question of Lorentz invariance is considered. We define a condition for the existence of elements of reality and a condition for the Lorentz invariance of these elements of reality. Then we show that, by considering a particular gedanken experiment, we obtain a contradiction demonstrating that Lorentz-invariant realistic interpretations of quantum theory are not possible.

Acknowledgements

I have enjoyed and benefited from conversations with many people during the course of this work. It gives me great pleasure to thank my supervisor, Euan Squires, for always being available to discuss new ideas and for his continued help and encouragement over the last three years. I would like to thank audiences at Cambridge, Innsbruck, Oxford, Portsmouth, and Sussex and also the many people I met at conferences in Goslar and Cesena, in particular, Gino Lepore, Pan Kaloyerou, Marek Żukowski and Anton Zeilinger. I would like to thank Chris Dewdney and Mai Mai Lam for helping me to understand the de Broglie-Bohm model, Rob Clifton and Constantine Pagonis for discussions during my visits to Cambridge and their visits to Durham, and Michael Lockwood for a conversation on Lorentz invariance in Orkney. In addition, I would like to express my thanks to Partha Ghose and Dipankar Home for discussions during their visits to Durham and to David Hind for many conversations on conceptual matters.

This work has been funded by the Science and Engineering Research Council to whom I am grateful.

I would like to thank my friends in Durham for making my time here enjoyable. Finally, I would like to thank my family who have helped me in more ways than I could possibly list here.

Contents

Chapter 1

Introduction

Although quantum mechanics is now about fifty years old, it is as mysterious if not more mysterious now than when it was first discovered. Whilst the application of the quantum mechanical formalism is relatively straight forward and furthermore, its predictions have so far always been verified, there is no straight forward way of understanding what the theory tells us about reality. There are perhaps four features of quantum theory which can be identified as being the root cause of the interpretational problems.

(1) The superposition principle.

(2) The projection postulate.

(3) The Heisenberg uncertainty principle.

(4) The fact that the quantum mechanical state must be written as a vector in the product space of the subsystems.

We will consider these in turn below. To illustrate the following remarks consider a system with one degree of freedom on which we can measure the physical property Q. The corresponding operator is $\widehat{Q}$. As the system has only one degree of freedom there are no other physical quantities that can be measured on this system with corresponding operators that commute with $\widehat{Q}$. The eigenvalue equation is $\widehat{Q}|q_i\rangle = q_i|q_i\rangle$. The state of the system $|\psi\rangle$ can be written as a superposition of the various states available to it, i.e.

$$|\psi\rangle = \sum_i c_i|q_i\rangle \tag{1.1}$$

where c_i are complex numbers which can be regarded as the probability amplitude associated with the states $|q_i\rangle$. The probability of that a measurement of Q will give the value q_i is equal to $|c_i|^2$. The total probability must be equal to one and consequently we have

$$\sum_i |c_i|^2 = 1 \tag{1.2}$$

The superposition principle states we add the probability amplitudes c_i and not the probabilities $|c_i|^2$ associated with each route by which the system could have evolved to the $|q_i\rangle$ state and then this sum is squared to calculate the probability that a measurement will reveal that the system is in this state. It is this which gives rise to interference effects. In classical mechanics we add probabilities because it is assumed that an individual system is always in actually in one state and that it evolves to this state by one route although other routes may be available and furthermore that the existence of the routes not taken does not influence the evolution of the individual system. The fact that we cannot do this in quantum mechanics calls into question these assumptions in the case of quantum systems.

The projection postulate says that after a measurement of the physical quantity Q yielding the value q_i, the state of the system becomes $|q_i\rangle$. This is disturbing because it requires that there are two types of evolution in quantum theory. The smooth unitary evolution when the system is not disturbed by measurements and the sudden 'jumps' or 'collapses' when the system undergoes a measurement. Furthermore, it is not clear exactly what constitutes a measurement and at what stage a measurement has been completed. This last point is particularly serious because, in principle, one could expect interference between two macroscopic states of a measurement apparatus if we assume that there is still a superposition. This would lead to different empirical predictions to those obtained by assuming that the system has collapsed onto one of the two states. In practice, one could not actually distinguish these two sets of predictions. However, if we are interested in the nature of reality then we cannot ignore a problem because it is not a problem in practice.

The Heisenberg uncertainity principle says that if Q and P are two physical quantities whose corresponding operators do not commute then the more precisely we can predict what the outcome of a measurement of one of these quantities would be the less precisely we can predict what the outcome of a measurement of the other quantity would be. Thus, if we can predict exactly what the outcome of a measurement of one quantity would be then we cannot make any prediction for what the outcome of a measurement of the other quantity would be. Consequently, there is an irremovable element of unpredictability in quantum theory. If the quantum mechanical state vector is to be regarded as a complete description of a system then we would have to say that quantum mechanics is in fact indeterministic. Determinism can always be restored by admitting the existence of hidden variables. Nevertheless, whether quantum mechanics should be given a deterministic interpretation remains a point of debate.

The need to write the state of a system as a vector in the product space of the subsystems is responsible for the properties of nonlocality and non-Lorentz invariance with which we shall be concerned in later chapters. Consider two subsystems 1 and

2 each with one degree of freedom. Let A and B be physical quantities that can be measured on each of these subsystems 1 and 2 respectively. The eigenvalue equations are $\widehat{A}|a_i\rangle = a_i|a_i\rangle$ and $B|b_j\rangle = b_j|b_j\rangle$. The state of the combined system is in general written as

$$|\Psi\rangle = \sum_{i,j} c_{ij}|a_i\rangle|b_j\rangle \tag{1.3}$$

Before the two subsystems have interacted we will be able to write $c_{ij} = d_i e_j$, i.e. the state of the system can be written as the product

$$|\Psi\rangle = \Big(\sum_i d_i|a_i\rangle\Big)\Big(\sum_j e_j|b_j\rangle\Big) \tag{1.4}$$

Thus the two subsystems can be regarded to be independent. However, after the two systems have interacted, it will no longer be possible to write their state as a product and we can no longer regard the two systems as being separate even though the interaction is over. This is different to classical mechanics as in classical mechanics the state of each subsystem can always be described independently of the other subsystem both before and after any interaction.

Historically, the interpretation of quantum mechanics has been a debate between the positivist motivated Copenhagen school and the realist tradition. In its strongest form, positivism states that it is only properties that are observed that can be considered real. If pushed, this could lead to a kind of solipsism in which one only regards one's own thoughts as real. The Copenhagen school believes in a weaker form of positivism in which only classical properties are considered in discussions about reality. Questions like 'where is the electron?' are met with the answer that this question is not a valid question except in the case when the position of the electron has been measured such that there is a corresponding classical property, namely the apparatus reading. There are substantial problems with this approach. Firstly, it is not clear where the dividing line between classical properties and quantum properties should be drawn and secondly, no explanation is given as to why certain questions are not valid. The realist approach attempts to answer questions like 'where is the electron?' or at least give a reason as to why the question is not valid. The de Broglie-Bohm model is the best example of a realist interpretation. In this interpretation the particles always have an exact position and momentum and furthermore there is no dividing line between the quantum and the classical world. We will find, however, that the price to be paid for realism is nonlocality and non-Lorentz invariance.

In this work we will address some of the problems of realist interpretations. Chapters 2 and 3 concern wave-particle duality. If a single particle state impinges onto a beam splitter and a detector is placed in each output then only one of the two detectors will detect the particle. This is due to the projection postulate and can be seen as an illustration of the particle nature of quantum systems. Now, if the detectors are removed and the two paths are recombined at a second beam splitter then interference will be observed. This is due to the superposition principle. In chapter 2 we will consider such single particle quantum phenomena. After seeing that a particle picture alone cannot explain all single particle quantum phenomena we consider the possibility of explaining such phenomena by classical wave theory. This attempt does, of course, fail but the exact conditions under which it fails have not been well tested in experiments. Rather than attempting to explain quantum mechanics purely in terms of waves or purely in terms of particles we could explain it in terms of waves and particles, both being taken to exist objectively (as in the de Broglie-Bohm model for example). If we take this approach then when a single particle impinges on a beam splitter the particle will only go one way but the wave will be split and go along both paths. The part of the wave that goes along the other path to the particle will be empty, that is it will not have a particle in it. It has been a point of much debate as to whether these empty waves can be taken to be real. In chapter 3 we will show how it is possible to demonstrate the reality of these empty waves if we allow ourselves to make three realist motivated assumptions.

Chapters 4 to 8 concern nonlocality in quantum mechanics. In chapter 4 we review the Einstein, Podolsky, and Rosen argument for the incompleteness of quantum mechanics and Bell's demonstration of nonlocality in realistic interpretations of quantum theory which we illustrate with the Clauser, Horne, Shimony, and Holt inequalities. Then we present a new way to obtain Bell inequalities that brings out more clearly the contradiction between quantum mechanics and local realism. In chapter 5 we review the use of supplementary assumptions in deriving Bell-type inequalities that can be tested in real experiments with existing technology. Then we present some new ways of obtaining correlated states that can be used to illustrate Bell's theorem. In chapter 6 we consider a recent proposal of Tan, Walls, Collett to demonstrate the nonlocality of a single photon. We find that their demonstration requires that an assumption is made that cannot be tested. This raises the question of whether the nonlocality of a single photon can be demonstrated without making any additional assumptions. We consider another proposal to demonstrate the nonlocality of a single photon. This demonstration does not require any additional assumptions but it has a rather more serious weakness if it is to be regarded as a demonstration of the nonlocality of a single photon. However we find that we can remove this weakness and the need for an additional assumption if we combine the two proposals. In chapter 7 we review the recent demonstration of nonlocality in quantum mechanics due to Greenberger, Horne, and Zeilinger (GHZ) which uses three or more particles. This demonstration does not require inequalities in the ideal case. However, in the non-ideal case, inequalities are required. We show how such inequalities can be obtained by using the methods used in chapter 4 to obtain two particle inequalities. We also present some ways of obtaining N-particle states that can be used to illustrate these arguments. The GHZ demonstration of nonlocality requires three or more particles. The GHZ approach cannot be applied to two particles to demonstrate nonlocality without using inequalities. In chapter 8 we present a new gedanken experiment that can be used to demonstrate nonlocality in quantum mechanics without using inequalities but that requires only two particles. Once again, in a real experiment inequalities will be required. We derive inequalities that are directly applicable to a quantum optical version of the experiment.

In chapter 9 we discuss the question of Lorentz invariance in quantum theory. A condition for the existence of elements of reality similar to that of Einstein, Podolsky, and Rosen is defined and also a condition for the Lorentz invariance of the elements of reality is defined. It is found that when these conditions are applied to the gedanken experiment considered in chapter 8 we obtain a contradiction. Consequently, quantum mechanics cannot be given a Lorentz invariant realistic interpretation. This contradiction can be illustrated in a most dramatic way if we assume that particles have real trajectories at least in those situations in which there is only one path open for the particle concerned to have reached a particular region and the particle is known to be in that region

Chapter 2

Wave-Particle Duality

2.1 Introduction

When a single particle wave packet impinges on a beam splitter, the particle can only be detected in one of the two output beams for a given run of the experiment. We shall refer to such experiments as beam splitter anticorrelation experiments. To explain such experiments, a particle picture of quantum phenomena will suffice. If the two output beams from the beam splitter are recombined at a second beam splitter (e.g. as in a Mach-Zehnder interferometer) then interference can be observed by varying the path difference between the two paths through the interferometer. To explain this interference a wave picture of light will suffice. Thus, it seems that the same quantum system will sometimes behave as a particle and sometimes behave as a wave.

We will consider a particle to be a concentration of matter occupying a small volume of space with the property that when there are two or more possible paths along which it could go it will only actually go along one. It is clear that interference phenomena such as that in a Mach-Zehnder interferometer cannot be explained in terms of particles alone because the particle will only transverse one path through the interferometer and therefore cannot pick up any information about the other path. Thus the behaviour of that particle cannot depend on the length of the other path. If the two paths through the interferometer are set equal then quantum theory predicts that there will be destructive interference in one of the two outputs. However, if the particle has no information about the other path then it has no way of 'knowing' that the path lengths have been set equal and consequently it may go into the dark output thus violating the predictions of quantum theory. However, as we shall see it is not so clear that beam splitter anticorrelation experiments cannot be explained in terms of waves alone. This is particularly true of photons because their discrete properties can often be explained in terms of the discrete properties of the detection process rather than in terms of the photon field itself having any discrete properties (see Loudon (1980)). In this chapter we shall show how the attempt to explain the photon anticorrelation effect at a beam splitter in terms of classical wave theory fails. In this chapter we shall only be concerned with single photon phenomena.

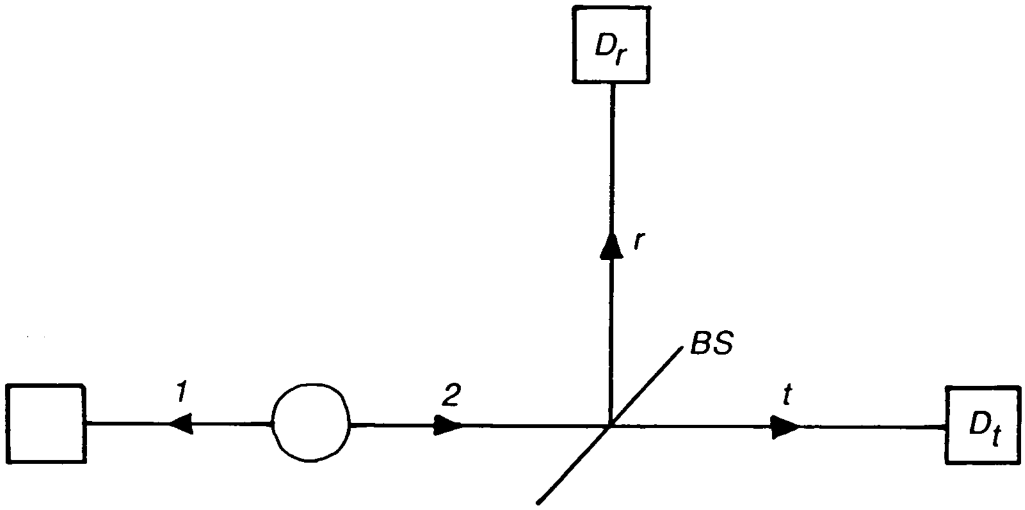

Fig. 2.1. Pairs of photons are produced by an atomic cascade. Detection of photon 1 activates the detectors in the reflected and transmitted outputs from the beam splitter.

However, two photon experiments have been performed which also address the question of whether classical wave theory can explain quantum phenomena, for example Franson (1991) and the experiments to test local realism (which we will discuss in chapter 5).

2.2 The experiments of Grangier et al.

Using a source of nearly ideal single photons, Grangier, Roger and Aspect (1986) performed both a beam splitter anticorrelation experiment and an interference experiment providing experimental evidence for quantum theory in this domain where wave-particle duality is most manifest (see also Grangier (1986), Aspect and Grangier (1987) and Aspect (1990)). To prepare a source of nearly ideal single photons they used an atomic cascade which produces pairs of photons. When one of these photons is detected then we know that we have another photon. Thus the first photon can be used to turn on the detectors for some short time ω (typically 10ns) ready for the second photon. The anticorrelation experiment they performed is shown in fig. 2.1. Photon pairs each consisting of photon 1 and photon 2 are produced in an atomic cascade source. Detection of photon 1 produces a gate of duration ω during which the detectors D_r and D_t are active. The detectors D_r and D_t are placed in the reflected and the transmitted outputs of the beam splitter BS respectively.

Grangier et al. considered how this experiment would be described by classical wave theory. During the nth gate the probability of a count is, according to classical wave theory,

$$p'_n = \alpha' \int_{t_n}^{t_n+\omega} I'_n(t)dt \tag{2.1}$$

where $I'_n(t)$ is the intensity incident on the detector and α' is a constant characteristic of the detector. Averaged over an ensemble of N gates the probability of a count is

$$p' = \alpha' \omega \langle I'_n \rangle \tag{2.2}$$

where

$$\langle I'_n \rangle = \frac{1}{N} \sum_{n=1}^{N} \frac{1}{\omega} \int_{t_n}^{t_n+\omega} I'_n(t)dt \tag{2.3}$$

Let p_r (p_t) be detection probability in the reflected (transmitted) beam, that is the number of detections at detector D_r (D_t) divided by the total number of gates when the total number of gates is large. Let p_c be the coincidence probability, that is the number of times there is a detection at both D_r and D_t during the same gate divided by the total number of gates when the total number of gates is large. If the intensity incident on the beam splitter during the nth gate is I_n then the reflected (transmitted) intensity is RI_n (TI_n) where R (T) is the reflectance (transmissivity). Thus, using equation (2.1) we have

$$\begin{aligned} p_r &= \alpha_r \omega R \langle I_n \rangle \\ p_t &= \alpha_t \omega T \langle I_n \rangle \\ p_c &= \alpha_r \alpha_t \omega^2 RT \langle I_n^2 \rangle \end{aligned} \tag{2.4}$$

The Cauchy-Schwarz inequality is

$$\langle I_n^2 \rangle \geq \langle I_n \rangle^2 \tag{2.5}$$

Using this and equations (2.4) we obtain

$$\alpha \geq 1 \tag{2.6}$$

where

$$\alpha = \frac{p_c}{p_r p_t} \tag{2.7}$$

Inequality (2.6) was derived and tested by Grangier et al. They obtained a value of $\alpha = 0.18 \pm 0.06$ which violates the inequality. It would appear that classical wave theory is not capable of explaining this experimental result and so the particle picture is required.

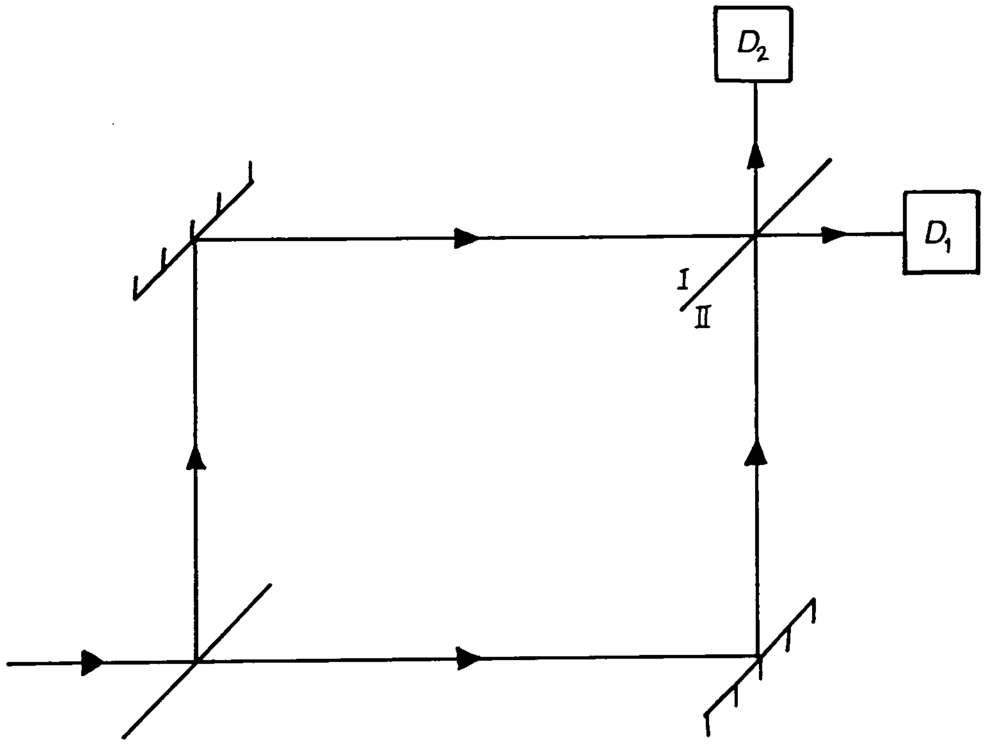

Fig. 2.2. Interference experiment performed by Grangier et. al.

Grangier et al. also performed an interference experiment shown in fig. 2.2 using the same source as in the anticorrelation experiment to demonstrate the wave behaviour of light. The reflected and transmitted beams were recombined using a second beam splitter giving fringes with visibility $V \approx 0.98$.

2.3 Stochastic transmissivity

It was pointed out by Marshal and Santos (1987) that the derivation of inequality (2.6) requires that the transmissivity, T, and the reflectance, R, are constants. If we assume, instead, that they are stochastic variables then α can take values less than 1 according to wave theory. Thus classical wave theory could explain the anticorrelation experiment. In the extreme case where $T(t)$ and $R(t)$ take the values 0 and 1 only, an individual wave packet may be entirely reflected or entirely transmitted. This would give $\alpha = 0$ (zero probability of coincidence). However, if we assume that $T(t)$ and $R(t)$ have the same distribution in an interference experiment then, for this case, there would be no interference (i.e. $V = 0$) as the amplitude in one of the two beams would always be zero. In a less extreme case where $T(t)$ and $R(t)$ take values from 0 to 1 we may still have $\alpha \leq 1$

with some interference effect (see Hardy (1991a)). Clearly, what is required is a way of comparing the two experiments to see if classical wave theory with stochastic variation of the transmissivity and reflectance can explain the results of both experiments.

For the nth gate let the incident beam have amplitude

$$A_n(x,t) = a_n(x,t)e^{ik(x-ct)} \tag{2.8}$$

where $a(x,t)$ is an envelope function. The reflected and transmitted amplitudes are given by

$$A_n^r(x_r,t) = \sqrt{R(t-x_r/c)}A_n(x_r,t) \tag{2.9}$$

$$A_n^t(x_t,t) = \sqrt{T(t-x_t/c)}A_n(x_t,t) \tag{2.10}$$

where the beam splitter is placed at $x = 0$ and x_r (x_t) is the distance measured along the reflected (transmitted) beam. Using mirrors these two beams can be superposed at $x = x_r = x_t + \delta$ (note that here we are not using a second beam splitter to recombine the beams). Assuming $a(x,t)$ varies on a distance scale much greater than δ and $T(t)$ and $R(t)$ vary on a time scale much longer than δ/c (this assumption is certainly justified if we only consider the first few interference fringes where $|\delta| \leq 2\lambda$) we get the resulting superposition intensity

$$I_n^s(t) = R(t')I_n(t) + T(t')I_n(t) + 2\sqrt{R(t')T(t')}I_n(t)\cos(k\delta) \tag{2.11}$$

where $t' = t - x/c$ and $I_n(t) = |a_n(x,t)|^2$ is the incident beam intensity. The detection probability is

$$p_s = \alpha_s\omega\langle I_n^s\rangle \tag{2.12}$$

The fringe visibility, $V_s = (p_s^{max} - p_s^{min})/(p_s^{max} + p_s^{min})$, is then

$$V_s = \frac{2\langle\sqrt{RT}I_n\rangle}{\langle RI_n\rangle + \langle TI_n\rangle} \tag{2.13}$$

In the anticorrelation experiment

$$p_r = \alpha_r\omega\langle RI_n\rangle$$

$$p_t = \alpha_t\omega\langle TI_n\rangle \tag{2.14}$$

$$p_c = \alpha_r\alpha_t\omega^2\langle RTI_n^2\rangle$$

p_r should also be measured with the other detector, i.e.

$$p_r' = \alpha_t \omega \langle RI_n \rangle \tag{2.15}$$

Now $\alpha = p_c/(p_r p_t)$. Hence

$$\alpha = \frac{\langle RTI_n^2 \rangle}{\langle RI_n \rangle \langle TI_n \rangle} \tag{2.16}$$

The Cauchy-Schwarz inequality gives

$$\langle RTI_n^2 \rangle \geq \langle \sqrt{RT} I_n \rangle^2 \tag{2.17}$$

If we put

$$R_0 = \frac{p_r'}{p_r' + p_t} \tag{2.18}$$

$$T_0 = \frac{p_t}{p_r' + p_t} \tag{2.19}$$

then using (2.13) to (2.19) we get

$$\alpha \geq \frac{V_s^2}{4R_0 T_0} \tag{2.20}$$

This inequality now replaces inequality (2.6). Note that if $V_s = 1$ and $T_0 = R_0 = \frac{1}{2}$ then (2.20) becomes identical with (2.6).

From inequality (2.20) we see that if we wish to show that classical wave theory cannot explain the anticorrelation experiment then we must appeal not only to the results of that experiment but also to the results of another experiment – the interference experiment – which uses the same source of light and the same beam splitter. It is rather ironic that we must appeal to an interference experiment which demonstrates the wave behaviour of light to show that the wave picture alone is inadequate.

2.4 Do performed experiments violate clasical Wave theory?

Unfortunately the interference experiment of Grangier et al. was not quite of the form considered above. Instead a second beam splitter was used to recombine the reflected and transmitted beams from the first beam splitter. From the second beam splitter two more beams emerge. By varying the path difference interference fringes can be observed in each of these two beams with visibility V_1 and V_2. Let the second beam splitter have reflectance (transmissivity) R' (T') for light incident on one side (side I, see fig 2.2) and R'' (T'') for light incident on the other side (side II, see fig 2.2). Beam one is the result of superposing light which is reflected at the first beam splitter and transmitted at the second beam splitter (being incident on side I) with light which is transmitted at the first beam splitter and reflected off side II of the second beam splitter. Thus replacing R with RT' and T with TR'' in (2.13) gives V_1:

$$V_1 = \frac{2\langle\sqrt{RT'TR''}I_n\rangle}{\langle RT'I_n\rangle + \langle TR''I_n\rangle} \tag{2.21}$$

V_2 is obtained similarly by replacing R with RR' and T with TT'' in (2.13):

$$V_2 = \frac{2\langle\sqrt{RR'TT''}I_n\rangle}{\langle RR'I_n\rangle + \langle TT''I_n\rangle} \tag{2.22}$$

We define V by

$$\frac{2}{V} = \frac{1}{V_1} + \frac{1}{V_2} \tag{2.23}$$

By making some assumptions we can show that

$$V_S \geq V \tag{2.24}$$

These assumptions are:

(i) The stochastic processes at the source and at each of the two beam splitters are uncorrelated, i.e.

$$\langle f(R,T)g(R',T',R'',T'')h(I_n)\rangle = \langle f(R,T)\rangle\langle g(R',T',R'',T'')\rangle\langle h(I_n)\rangle \tag{2.25}$$

where f, g and h are any functions.

(ii) The symmetry property

$$\langle g(R',T',R'',T'')\rangle = \langle g(R'',T'',R',T')\rangle \tag{2.26}$$

holds where g is any function.

From (2.26) we get

$$\langle\sqrt{R'T''}\rangle = \langle\sqrt{R''T'}\rangle \tag{2.27}$$

$$\langle R'\rangle = \langle R''\rangle \tag{2.28}$$

$$\langle T'\rangle = \langle T''\rangle \tag{2.29}$$

Using (2.21) to (2.23), (2.25) and (2.27) to (2.29) we find

$$\frac{2}{V} = \frac{\langle R\rangle + \langle T\rangle}{2\langle\sqrt{RT}\rangle}\cdot\frac{\langle R'\rangle + \langle T''\rangle}{\langle\sqrt{R'T''}\rangle} \tag{2.30}$$

But (2.13) and (2.25) give

$$V_s = \frac{2\langle\sqrt{RT}\rangle}{\langle R\rangle + \langle T\rangle} \tag{2.31}$$

Hence

$$V = kV_s \tag{2.32}$$

where

$$k = \frac{2\langle\sqrt{R'T''}\rangle}{\langle R'\rangle + \langle T''\rangle} \tag{2.33}$$

Since

$$\langle(\sqrt{R'} - \sqrt{T''})^2\rangle \geq 0$$

we obtain

$$k \leq 1 \tag{2.34}$$

(2.32) and (2.34) give (2.24). Substituting (2.24) in (2.20) we get

$$\alpha \geq \frac{V^2}{4R_0T_0} \tag{2.35}$$

Grangier et al. measured $\alpha = 0.18 \pm 0.06$, $V \approx 0.98$ and $R_0 \approx T_0 \approx 0.5$. These values yield a significant violation of (2.35). This would suggest that the classical wave theory is inadequate if assumptions (i) and (ii) are true. However, only by direct measurement of V_s can we hope to finally rule out classical wave theory without the need for assumptions (i) and (ii). While the interference experiment with only one beam splitter (to measure V_s) would be more difficult to perform than the interference experiment with two beam splitters there is no reason why it cannot be performed at least to put a lower limit on V_s. In the above, classical wave theory is effectively defined by equation (2.1). A more sophisticated wave theory has been considered by Marshall and Santos (1988) in which the vacuum plays a crucial role and in which there is a threshold intensity below which a detector will not register a count. This theory, which Marshall and Santos call stochastic optics, is able to explain the results of experiments of Grangier et al. However, it seems a fair criticism that stochastic optics is a rather ad hoc theory in which the variable parameters are chosen after the experiment to fit the experimental data. Furthermore, as it is a local theory and therefore, it is clear that if the experimental conditions to test local realism are ever truly realised (i.e. without the need for supplementary assumptions) then stochastic optics is likely to be ruled out. However, as we will see in chapter 5, these conditions have not yet been realised and therefore, as Marshall and Santos (1989) have shown, the results of all experiments performed so far to test local realism can be accounted for by stochastic optics. Bohm's (1952) causal interpretation for photons does not treat photons as particles (although fermions are treated as particles) and therefore this might also be regarded as a wave theory. However, it is certainly not a classical wave theory because it is explicitly nonlocal and agrees with the predictions of quantum optics.

Chapter 3

Do Empty Waves Exist?

3.1 Introduction

In the previous chapter we saw that single particle quantum phenomena cannot be explained in terms of particles alone or in terms of waves alone. This is the origin of the so-called wave-particle duality of quantum theory (see Scully, Englert and Walther (1991), Ghose, Home and Agarwal (1991) and Ghose and Roy (1991)). One way out of this dilemma is to suppose that the particle is accompanied by a wave. Both the particle and the wave are assumed to exist objectively. When the particle plus wave impinge onto a beam splitter the particle goes one way but the wave is divided and goes in both directions. The wave that goes in the other direction to the particle will be empty, that is, it is not accompanied by a particle. If the two beams are recombined then the empty wave is recombined with the non-empty wave and the two can interfere. The wave could then serve to direct the particle in such a way that the probability density of finding the particle at some position is equal to the square modulus of the amplitude of the wave. This is the basic idea behind the de Broglie-Bohm interpretation of quantum mechanics (see Bohm (1952)). There have been many proposals for experiments to demonstrate the existence of these empty waves. See for example, Selleri (1969), Garuccio, Popper and Vigier (1981), Garuccio, Rapisarda and Vigier (1982), Tarozzi (1985), Croca (1987) and Selleri (1989). However, all of these experiments can be regarded as a test between some theory that explicitly proposes the existence of empty waves and quantum theory. They all have in common the fact that quantum mechanics predicts that the empty wave will not be detected in the way suggested. The purpose of this chapter is to show how we can demonstrate the reality of empty waves within quantum theory.

3.2 Conditions for the existence of empty waves

To demonstrate that the empty waves really exist, it is not sufficient to consider only standard interference experiments like the Mach-Zender interferometer. These experiments only illustrate the theoretical usefulness of the empty wave concept. In these experiments, the presence of the empty wave is only felt when it is recombined with the non-empty wave, i.e. when it is no longer empty. Furthermore, we cannot demonstrate that the particle only goes along one path but rather we must suppose that it does and even then we do not know which path this is. What is required is that the empty wave can manifest its reality in a region where we can establish that it is actually empty. This suggests the following *sufficient conditions for the existence of empty waves*:

(1) We know which path the particle goes along.

(2) Some measurable property of a system placed in another path (along which we suppose an empty wave goes) is changed.

If we can find a situation in which both these conditions are satisfied then we will have demonstrated the reality of empty waves. If there is only one path open for a particle to have reached a detector and that detector registers a detection then we cannot necessarily conclude that there was a particle that actually went along that path because we do not observe a particle going along that path. However, if we are to discuss empty waves then we must have some notion of particle trajectory at least in those situations where there is only one possible path between the source and the detector. Hence, if we are to proceed then we must make the following assumption

(i) If a there is only one path open for a particle to have reached a detector and that detector registers a detection then there was a particle that actually took that path.

If we make this assumption then there will be some situations in which we can satisfy the condition (1) above.

3.3 A naive way of trying to detect empty waves

A naive way of trying to satisfy conditions (1) and (2) to demonstrate the reality of empty waves is shown in fig. 3.1. A single particle impinges on a beam splitter BS. In output v a detector D is placed. In output u some system A is placed. If the particle is detected at detector D then, using assumption (i) above, we can deduce that the particle took path 1. Thus condition (1) above is satisfied. To satisfy condition (2) we require that the system A has some measurable property changed, at least for some runs of the experiment in which the particle is detected at D. However, we can show that regardless of the nature of the system A quantum mechanics predicts that condition (2) will not be satisfied (see Gordon (1989), Holland (1992) and Mandel (1984)). The evolution of the particle, initially in the input mode s, as it passes through the beam splitter is given by

$$|s\rangle \longrightarrow \frac{1}{\sqrt{2}}\Big(|u\rangle + i|v\rangle\Big) \tag{3.1}$$

Let $|A\rangle$ be the state of system A in the absence of interaction with any other systems and let $|A'\rangle$ be the state of the system after interaction with the $|u\rangle$ state. Thus we have

$$|u\rangle|A\rangle \longrightarrow |A'\rangle|u\rangle \tag{3.2}$$

Let $|D_0\rangle$ be the state of the detector when it has not fired and let $|D_1\rangle$ be its state when it has fired. The initial state of the whole system is

$$|s\rangle|A\rangle|D_0\rangle \tag{3.3}$$

After the particle passes through the beam splitter the state becomes

$$\frac{1}{\sqrt{2}}\Big(|u\rangle + i|v\rangle\Big)|A\rangle|D_0\rangle \tag{3.4}$$

Finally, after interaction with the detector D and system A, the state becomes

$$\frac{1}{\sqrt{2}}\Big(|u\rangle|A'\rangle|D_0\rangle + i|v\rangle|A\rangle|D_1\rangle\Big) \tag{3.5}$$

If the detector fires so that condition (1) is satisfied then the state of the system is projected on to the second term in (3.5). Thus the state of system A remains $|A\rangle$ and therefore there can be no change in any of its measurable properties. Thus condition (2) cannot be satisfied. Therefore we see that this set up cannot be used to demonstrate the reality of empty waves.

On the basis of this fact one might assume that empty waves can never manifest their reality if quantum theory is correct. However, we will see that if we allow ourselves to make two more very natural assumptions then we can demonstrate the reality of empty waves (that manifest their reality). This demonstration is possible because, in the gedanken experiment to be considered, we will be able to know with certainty that the particle went along one of two possible paths even though there is no direct measurement to confirm this (i.e. a detector is not placed in the path to detect the particle).

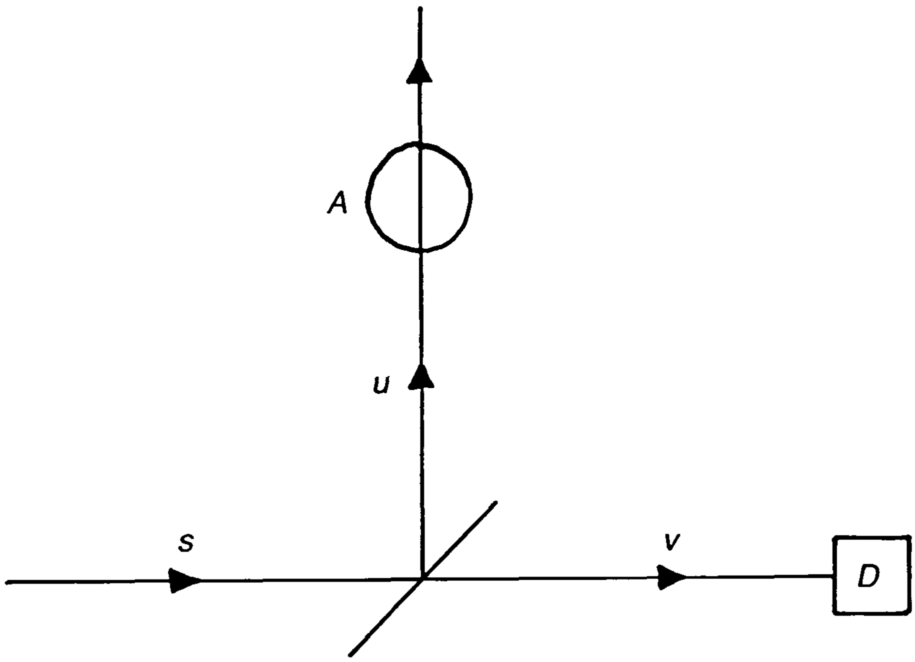

Fig. 3.1. If the particle is detected at D then we might naively expect it to be possible for system A to be affected.

3.4 The gedanken experiment of Elitzur and Vaidman

Before discussing empty waves we will consider a gedanken experiment due to Elitzur and Vaidman (1991). They considered a Mach-Zehnder type interferometer with a single particle source arranged so that, due to destructive interference, no particles will be detected at one of the two outputs. We will call this output the 'dark output' even though it will not be 'dark' in all of the situations we are going to consider. Now, if an object is placed in one of the two paths through the interferometer so as to block particles travelling along that path then it is possible for a particle to be detected at the dark output because there is no longer any destructive interference. We know that the particle must have taken the path without the object in it because otherwise it would have been absorbed by the object. In other words: If a particle is detected in the dark output then we know (1) that there must be an object in one path and (2) that the particle took the other path. Thus, we can deduce the presence of an object even when no particle impinges on it. Elitzur and Vaidman then considered the possibility that the object is a quantum particle, an atom say, which can be located in one of two disjoint regions of space A and B. The experiment is arranged so that a particle travelling along one path will be blocked by the atom if the atom is in region A but not if the atom is in region B. Particles travelling along the other path will be unaffected by the atom. (At this stage the experiment becomes a gedanken experiment because in any real experiment an atom would not block all the particles as is assumed here.) The atom can be prepared in a superposition of the states $|A\rangle$ (atom in region A) and $|B\rangle$ (atom in region B). For example

$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big(|A\rangle + |B\rangle\Big) \tag{3.6}$$

If the particle is detected in the dark output then one path must be blocked and therefore the state of the atom must be $|A\rangle$. This was presented by Elitzur and Vaidman as a form of interaction-free measurement because the particle has been used to establish the state of the atom without interacting directly with it (had the particle interacted directly with the atom in that state then it would have been absorbed).

3.5 A gedanken experiment for empty waves

We will now extend the gedanken experiment of Elitzur and Vaidman described above in order to demonstrate the reality of empty waves. Fig. 3.2 shows a Mach-Zehnder type interferometer for single particles. The state of a particle incident on the input is $|s\rangle$. The states of a particle in the two paths inside the interferometer are $|u\rangle$ and $|v\rangle$ for paths u and v respectively and the states of a particle in the two outputs are $|c\rangle$ and $|d\rangle$ for outputs c and d respectively. Detectors are placed in each of the two outputs. As before, the interferometer is arranged so that it has a 'dark output'. That is, if neither of the paths through it are obstructed in any way then, due to destructive interference, no particles will be detected at one of the two outputs, the d output say. Now, a spin $\frac{1}{2}$ atom is prepared with spin $+\frac{1}{2}$ along the x axis. In terms of z spin states the initial state of the atom can be written

$$|\text{atom}\rangle = \frac{1}{\sqrt{2}}\Big(|+\rangle + |-\rangle\Big) \tag{3.7}$$

where $|\pm\rangle$ is the state of the atom when it has spin $\pm\frac{1}{2}$ along the z axis.

The atom, prepared in this state, is placed in a closed box that is longer than it is wide (see fig. 3.3(a)). A nonuniform magnetic field is placed along the length of the box which is aligned along the z axis. This acts like a Stern-Gerlach apparatus and has the effect of splitting the state of the atom between the upper part of the box with the $|+\rangle$ state in it and the lower part of the box with the $|-\rangle$ in it. Next, dividing walls are placed so as to divide the box into upper and lower halves and these are then separated to form

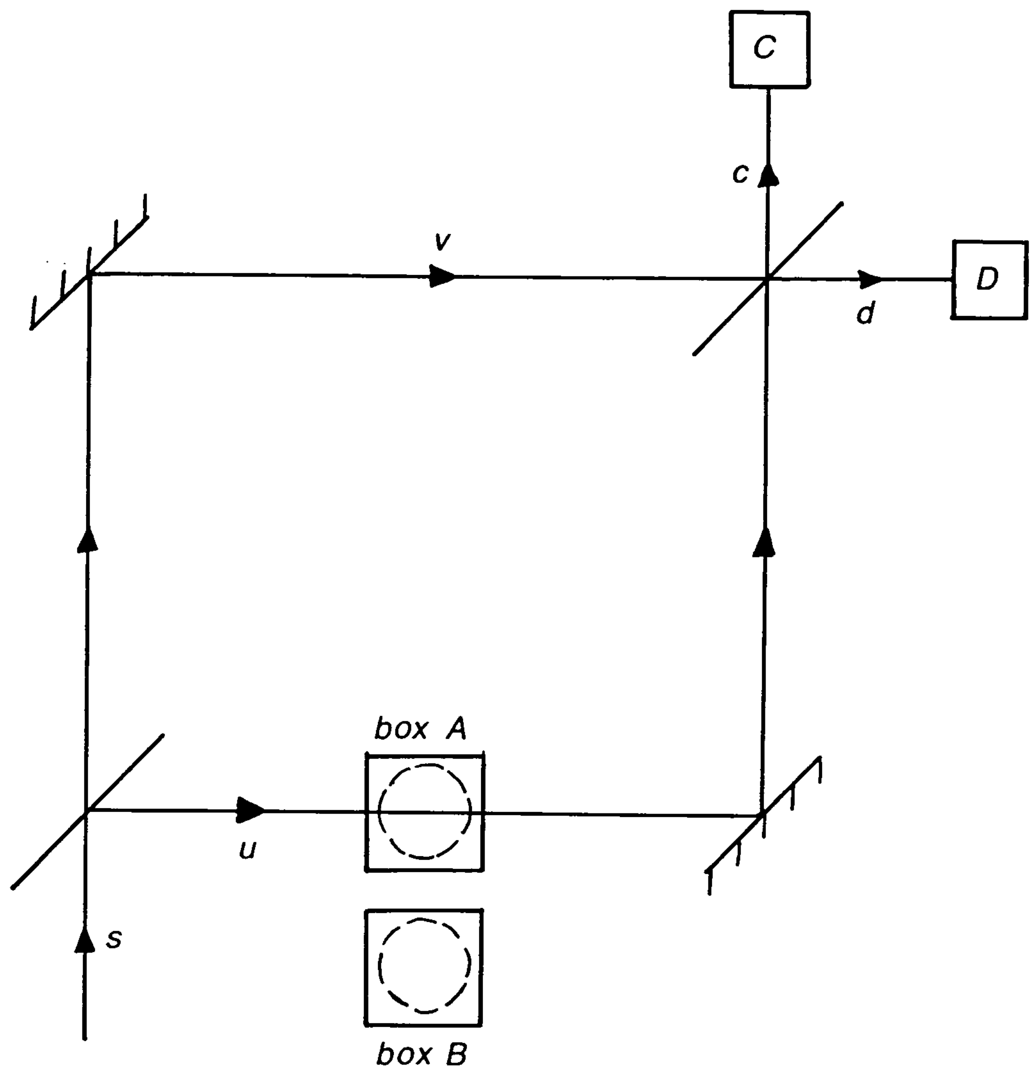

Fig. 3.2. Interferometer arranged such that, when there are no obstructions in either of the two paths, no particles will be detected in output d. Box A is then placed in path u.

two closed boxes (see fig. 3.3(b).) then the nonuniform magnetic field is removed. The upper box will be called box A and the lower box will be called box B. The state of this system is

$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big(|A,+\rangle + |B,-\rangle\Big) \tag{3.8}$$

where $|A,\pm\rangle$ ($|B,\pm\rangle$) is the state of the system when the atom is in box A (B) and has spin $\pm\frac{1}{2}$. The boxes are constructed from a material which is transparent to the interferometer particles but which cannot be penetrated by the atom. That is, the boxes

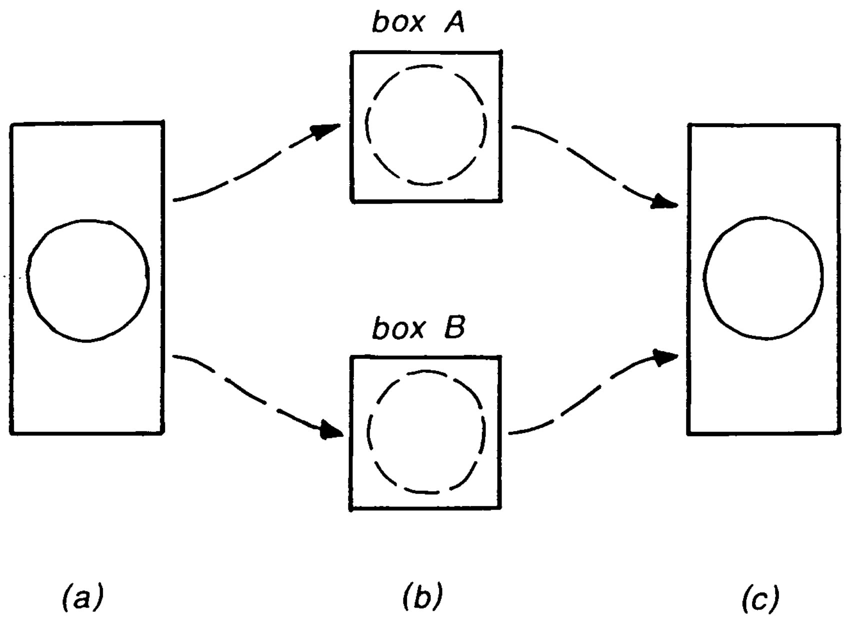

Fig. 3.3. (a) The atom, with spin $+\frac{1}{2}$ along the x axis is placed in a box and a non-uniform magnetic field is placed along the z axis. (b) Dividing walls are placed so as to create two boxes A and B. (c) The boxes are brought back together and the dividing walls are removed.

represent an infinite potential barrier to the atom but zero potential to the interferometer particles. Box A is placed in path u of the interferometer as shown in fig. 3.2. If the atom is in box A then we assume that it will absorb the particle with probability equal to one if the particle goes along path u. Expressed mathematically, this gives

$$|u\rangle|A,+\rangle \longrightarrow |A,+\rangle_{ex} \tag{3.9}$$

where $|A,+\rangle_{ex}$ is the state of the atom when it has absorbed the particle. After the particle has passed through the interferometer and been detected (we will not be interested in those cases when the particle is absorbed by the atom) the two boxes are brought back together and the dividing walls are removed (fig. 3.3(c).). Finally a measurement of spin along the x axis is made on the atom.

The operation of the first beam splitter on the particle is given by

$$|s\rangle \longrightarrow \frac{1}{\sqrt{2}}\Big(i|u\rangle + |v\rangle\Big) \tag{3.10}$$

The operation of the second beam splitter on the particle is given by

$$|u\rangle \longrightarrow \frac{1}{\sqrt{2}}\Big(|c\rangle + i|d\rangle\Big) \tag{3.11}$$

and

$$|v\rangle \longrightarrow \frac{1}{\sqrt{2}}\Big(i|c\rangle + |d\rangle\Big) \tag{3.12}$$

For the time the two boxes are separated the evolution of the whole system is given by, using (3.9) to (3.12),

$$\frac{1}{\sqrt{2}}|s\rangle\Big(|A,+\rangle + |B,-\rangle\Big) \tag{3.13}$$

$$\longrightarrow \frac{1}{2}\Big(i|u\rangle + |v\rangle\Big)\Big(|A,+\rangle + |B,-\rangle\Big) \tag{3.14}$$

$$\longrightarrow \frac{i}{2}|A,+\rangle_{ex} + \frac{1}{2}|v\rangle|A,+\rangle + \frac{1}{2}\Big(i|u\rangle + |v\rangle\Big)|B,-\rangle \tag{3.15}$$

$$\longrightarrow \frac{i}{2}|A,+\rangle_{ex} + \frac{i}{2\sqrt{2}}|c\rangle|A,+\rangle + \frac{1}{2\sqrt{2}}|d\rangle|A,+\rangle + \frac{i}{\sqrt{2}}|c\rangle|B,-\rangle \tag{3.16}$$

Before proceeding we will state the remaining assumptions. The second assumption is:

(ii) If, at time t, the state of the atom is $|A,+\rangle$ then the atom is actually in box A even if we do not make a measurement to establish this.

The atom cannot pass through the walls of the box either to get in to the box or to get out of it. This motivates the third assumption:

(iii) If the box is closed at time t_1 and we establish at some later time, t_2, that the atom is actually in the box then the atom is actually in the box for all times from t_1 till t_3 when the box is opened.

There is some similarity between this assumption and assumption (i). Both assume that if there is only one path open to the particle concerned to reach a particular region of space-time and the particle is known to be in that region then the particle must have taken that path. Both assumptions can be taken to be different applications of this one assumption. However, for our present purposes it is clearer to treat them separately. We are only interested in those runs of the experiment for which there is a detection in the dark output, d. When this happens the system is projected onto the third term in (3.16) such that the state of the atom becomes $|A,+\rangle$ until the box A is brought back together with box B. If the particle is detected in the dark output at time t_2 then, using (ii), we establish that at this time and afterwards (until the boxes are brought back together) the atom is actually in box A. Using (iii) we can now establish that the atom is actually in box A from the time, t_1, the dividing walls are put in place till the time, t_3, box A is brought together with box B and the dividing walls removed. If the atom is in box A, however, then path u is blocked. Consequently, the particle can only have reached the detector by taking path v through the interferometer. Thus, using assumption (i), we establish that the particle actually took path v through the interferometer. Therefore, we have satisfied condition (1) above for demonstrating the reality of empty waves. We will now show how condition (2) can be satisfied. If the particle goes along path v then we must demonstrate that the atom plus box system which is placed in path u, along which we suppose the empty wave goes, has some measurable property changed. If the experiment is performed exactly as described above except that no particle is sent through the interferometer then when the boxes are brought together and the measurement of spin of the atom along the x axis is made then the result will be $+\frac{1}{2}$ because the atom is initially prepared with spin $+\frac{1}{2}$ along the x axis. This can be repeated as many times as the experimenter wishes. The result will always be $+\frac{1}{2}$. Spin along the x axis is a measurable property of the atom plus box system that will take the value $+\frac{1}{2}$ every time the measurement is made unless the system is affected by something. If we now perform the full experiment, sending a particle through the interferometer, and the particle is detected in the dark output then the state of the atom becomes $|A,+\rangle$, i.e. it has spin $+\frac{1}{2}$ along the z axis. Consequently, when boxes are brought together and the spin along the x axis is measured, there is a 50% probability that it will take the value $-\frac{1}{2}$. If we consider those runs of the experiment in which it does take this value then we see that there is a change in the value of the measurable property of the atom plus box system even though only the empty wave impinges on it. This satisfies condition (2). Therefore, if assumptions (i), (ii) and (iii) are true then we have demonstrated the reality of empty waves.

Strictly speaking, what we have actually demonstrated is that 'something' goes along the path that the particle does not go along. We have not demonstrated that it has wave properties and it might be better called an empty 'something'. The idea that these empty 'somethings' have wave properties comes from the use of empty waves to explain interference experiments. However, it is not possible to use standard interference experiments in conjunction with the assumptions (i), (ii) and (iii) to demonstrate the existence of empty waves or empty 'somethings' because we cannot satisfy conditions (1) and (2). However, we shall continue to call the empty 'somethings' empty waves, it being understood that their wave nature has not been established.

The empty wave is necessary for the change in the measurable property of the atom plus box system. This is clear because if box A is placed just outside the path of the empty wave then there is no effect. However, we cannot say that the empty wave alone *causes* the change in the measurable property. The effect also depends on the detection in the dark mode. If the particle really travels along path v when there is to be a detection in the dark mode then a detector placed at the last moment into path v just before the second beam splitter will detect the particle. However, if the particle is detected in path v, then the empty wave in path u will not effect the atom plus box system because the state of the whole system is projected on to the term $\frac{1}{2}|v\rangle(|A,+\rangle + |B,-\rangle)$ in (3.15). However, the decision to place the detector in path v and the measurement on the atom are made in two distinct regions that can be separated by a spacelike distance. Therefore, whether there is an effect on the atom or not depends nonlocally on the measurement that is made on the particle. If fact, we will see in chapter 8 this apparatus can be used to demonstrate Bell's theorem without using inequalities (in the experiment considered in chapter 8 the atom plus box system is replaced by a second interferometer, but the essential physics is the same). Pagonis (1992) has suggested that this nonlocal effect may in itself be sufficient to explain the change in the measurable property of the atom without having to invoke the existence of empty waves. To see that this is not true consider an ensemble of atom plus boxes systems all simultaneously taken through the same procedure as that described above but with only one of them having its A box placed in the u path. It is only the atom plus boxes system with box A in the u path that can undergo a change in its measurable property. Therefore, if we want to suppose that nonlocality alone is sufficient to explain the effect we have to explain how this nonlocal effect 'knows' only to pick out the atom plus boxes system which has box A in path u and leave all the other systems unchanged. Clearly there is no way of doing this unless we allot some significance to the path without the particle in it. That is we must suppose that something exists in the path without the particle. In our language this is an empty wave.

A peculiar feature of the above arguments is that they are frame dependent. In order to establish that the atom is in box A using assumption (ii), it is necessary that the interferometer particle is detected *before* the two boxes are brought back together. However, these two events can be separated by a spacelike distance. Thus, in this context, the notion of 'before' is frame dependent. In chapter 9 we will use this gedanken experiment to show that any realist interpretation of quantum mechanics must violate Lorentz invariance. For our present purposes, this ambiguity can be removed by bringing the boxes together in the forward light cone of the detection of the particle. This way, the particle is detected before the boxes are brought together in all possible frames of reference.

3.6 Description in the de Broglie-Bohm model

The de Broglie-Bohm model for fermions (see Bohm, Hiley and Kaloyerou ((1987)) satisfies assumptions (i), (ii) and (iii). It is instructive to see explicitly how the de Broglie-Bohm model treats this gedanken experiment. In this model the particle and the atom each have an actual position at all times which we will denote by x and X respectively. The wavefunction in position representation is given by

$$\Psi(y,Y) = \langle y,Y|\Psi\rangle \tag{3.17}$$

where $|y,Y\rangle = |y\rangle|Y\rangle$ and $|y\rangle$ ($|Y\rangle$)is the position eigenstate corresponding to position y (Y) of the particle (atom). The quantum potential is

$$Q(x,X) = \left[-\frac{\hbar^2}{2m}\frac{\nabla_y^2 R(y,Y) + \nabla_Y^2 R(y,Y)}{R(y,Y)}\right]_{y=x,Y=X} \tag{3.18}$$

where $R^2 = |\Psi|^2$. The quantum potential acts on the particles in such a way that, if the probability density $\rho(x,X) = |\Psi(x,X)|^2$ at some time t_0, then the probability density continues to be equal to $|\Psi(x,X)|^2$ as the system evolves. An example of how the de Broglie-Bohm model treats two particle systems is given in Lam and Dewdney (1990).

At the first beam splitter the particle will either go along the u path in which case we will put $x = x_u$ or along the v path in which case we will put $x = x_v$. Similarly, when the boxes are separated, the atom will either be in box A in which case we will put $X = X_A$ or it will be in box B in which case we will put $X = X_B$. We will now consider the four possible cases:

(a) $x = x_u$ and $X = X_A$. When expanded, the state in (3.14) has four terms. When this state is evaluated with $x = x_u$ and $X = X_A$ (i.e. $\Psi(x_u, X_A)$) then only the $\frac{i}{2}|u\rangle|A,+\rangle$ term is non-zero. During the interaction between the atom and the particle this term evolves to $\frac{i}{2}|A,+\rangle_{ex}$ in (3.15). The other terms in (3.15) remain equal to zero and therefore do not contribute to the quantum potential acting on either the particle or the atom during this interaction. Therefore the particle will actually be absorbed by the atom.

(b) $x = x_u$ and $X = X_B$. The particle will not be absorbed by the atom and so we can consider the quantum potential acting on the particle as it passes through the second beam splitter. During this time the state evolves from that in (3.15) to that in (3.16). However, the atom remains in box B and therefore, when the state is evaluated at $X = X_B$, only the $\frac{1}{2}\Big(i|u\rangle + |v\rangle\Big)|B,-\rangle$ term in (3.15) which evolves to the $\frac{i}{\sqrt{2}}|c\rangle|B,-\rangle$ term in (3.16) is non-zero and can contribute to the quantum potential acting on the particle. Consequently the effect of this quantum potential must be to direct the particle into the c output. (There is no $|d\rangle|B,-\rangle$ term in (3.16) and so the particle cannot go into the d output in these cases.)

(c) $x = x_v$ and $X = X_B$. This case is similar to case (b) discussed above because $X = X_B$ and consequently we have a similar conclusion; the quantum potential will direct the particle in to output c.

(d) $x = x_v$ and $X = X_A$. These are the cases in which the empty wave is detected. The atom remains in box A as the particle goes through the beam splitter and therefore, when the state is evaluated at $X = X_A$, only the $\frac{1}{2}|v\rangle|A,+\rangle$ in (3.15) which evolves to $\frac{i}{2\sqrt{2}}|c\rangle|A,+\rangle + \frac{1}{2\sqrt{2}}|d\rangle|A,+\rangle$ in (3.16) non-zero and can contribute to the quantum potential acting on the particle. Thus, the quantum potential will send the particle into the d output in half of these cases. If we consider a case where the particle goes into the d output ($x = x_d$) then, when (3.16) is evaluated with $x = x_d$, only the $\frac{1}{2\sqrt{2}}|d\rangle|A,+\rangle$ term is non-zero. Therefore, when the boxes A and B are brought back together, the $|-\rangle$ state will not contribute to the quantum potential acting on the atom and consequently, a measurement of spin of the atom along the x axis will give $-\frac{1}{2}$ in half of the cases.

It is only in case (d) that the particle can be detected in the dark output, d. In this case, the particle actually goes along path v thus satisfying condition (1) and also the measurement of spin along the x axis on the atom may give $-\frac{1}{2}$ satisfying condition (2). Therefore, this is the first situation in which it has been shown that de Broglie-Bohm waves can manifest their reality in a region where they are empty. However, the Bohm model for bosons (see Bohm et al. (1987)) does not satisfy (i) if the interferometer particle is a boson because, unlike the de Broglie-Bohm model for fermions, this model does not treat quanta as particles but as fields. Nevertheless, even with this model, the field is still blocked from going past box A in path u if the atom is that box and, therefore, can only reach the detector by path v. That is, an influence (i.e. the change in the measurement of spin along the x axis of the atom) is felt in path u even though the field reaches the detector by path v. It should be noted, however, that the above arguments demonstrating the reality of empty waves are general. They do not only apply to the de Broglie-Bohm model for fermions, but to any interpretation in which the assumptions (i), (ii) and (iii) are satisfied. In particular, we do not necessarily assume the existence of particle trajectories in those situations where there is more than one possible path for the particle to go along. Whilst all three of these assumptions are not necessary for a realist interpretation of quantum mechanics, (the Bohm model for bosons being a good example of a realist interpretation in which they are not all true) they are, nevertheless, very natural to the realist conception of nature. Furthermore, they are not inconsistent with quantum mechanics.

Chapter 4

Nonlocality in Quantum Theory

4.1 Introduction

In 1935 Einstein, Podolsky and Rosen wrote a paper claiming to prove that quantum mechanics is incomplete. In their argument they made an implicit assumption of locality, i.e. that the choice of measurement made on one system does not influence another space-like separated system. In 1952 Bohm wrote down an hidden-variable model of quantum mechanics thus accomplishing the task of completing the theory as Einstein, Podolsky and Rosen had required. Actually, Bohm rediscovered an approach originally due to de Broglie. The de Broglie-Bohm model (as it is now generally called) is explicitly nonlocal thus contradicting the assumption of locality in the EPR argument. Einstein rejected the model on the basis of this nonlocality. However, in 1964 Bell demonstrated that any hidden-variable interpretation of quantum mechanics must be nonlocal. Thus, the EPR argument is flawed because it assumes locality and Einstein's criticism of the de Broglie-Bohm model is invalid in the sense that any complete interpretation of quantum mechanics must be nonlocal. Bell demonstrated that hidden-variable interpretations must be nonlocal by using the assumption of locality to derive a set of inequalities and then showing that these inequalities can be violated by the predictions of quantum mechanics.

In this chapter, we will review the EPR argument. Then we will review Bell's 1964 theorem and subsequent work done by Clauser, Horne, Shimony, and Holt (1969). Finally a new approach to Bell's theorem which provides some additional insight into the phenomenon of quantum nonlocality will be presented. All the work on Bell's theorem in this chapter will involve two particles in a singlet type state and the use of inequalities. In later chapters we will discuss other approaches to Bell's theorem using more than two particles or using just two particles but not prepared in a singlet type state.

4.2 The Einstein, Podolsky, Rosen argument

The 1935 paper by Einstein Podolsky and Rosen questioned the completeness of quantum mechanics. To do this in a rational way a set of criteria were required to define what is meant by a 'complete theory' and what constitutes an 'element of reality'. In a complete theory "every element of the physical reality must have a counterpart in the physical theory." Their criterion for the existence of an element of physical reality is the following: "If, without in any way disturbing a system, we can predict with certainty, (i.e. with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity." This criterion is regarded by EPR as a sufficient but not as a necessary condition for the existence of an element of reality which is in agreement with both classical and quantum mechanical ideas about reality.

Einstein, Podolsky and Rosen then go on to interpret quantum mechanics in the context of the criterion for completeness. If the operators corresponding to two physical quantities, say $\widehat{A}$ and $\widehat{B}$, do not commute (i.e. $\widehat{A}\widehat{B} - \widehat{B}\widehat{A} \neq 0$) then, if we can predict with certainty the result of measuring one of the two quantities, then we it is impossible to predict what the result of measuring the other will be. That is, in quantum mechanics the values of two quantities corresponding to two non-commuting operators can not be simultaneously known exactly. From this it can be seen that if two such quantities can be shown to have simultaneous physical reality (according to the criterion for physical reality) then, according to the criterion for completeness, quantum mechanics would not be complete. As EPR put it: "For if both of them had simultaneous reality – and thus definite values – these values would enter into the complete description according to the condition of completeness." The state vector, which provides the most complete description of a system available in quantum theory, cannot contain both of these exact values. In most situations it is not possible to allot physical reality to two non-commuting observables. In their paper EPR outline a situation in which it is possible to do this (by making use of the reality criterion). However, as we will see, there is an assumption of locality implicit in their argument.

Consider two systems I and II which interact from time $t = 0$ till $t = T$. For times after T there is no interaction between the systems. The state of the total system will be written $|\Psi(x_1, x_2)\rangle$ where x_1 and x_2 are the coordinates of systems I and II respectively. Now, let the physical quantity A pertaining to system I have eigenvalues a_1, a_2, a_3 . . . and corresponding eigenstates $|u_1(x_1)\rangle$, $|u_2(x_1)\rangle$, $|u_3(x_1)\rangle$ . . . It is now possible to express $|\Psi\rangle$ as

$$|\Psi(x_1,x_2)\rangle = \sum_n |\psi_n(x_2)\rangle|u_n(x_1)\rangle \tag{4.1}$$

where $|\psi_n(x_2)\rangle$ pertains to system II. The eigenvalues and eigenstates of another physical quantity B also pertaining to system I can be represented by b_1, b_2, b_3 . . . and $|v_1(x_1)\rangle$, $|v_2(x_1)\rangle$, $|v_3(x_1)\rangle$ . . . such that $|\Psi(x_1,x_2)\rangle$ can also be written as

$$|\Psi(x_1,x_2)\rangle = \sum_n |\varphi_n(x_2)\rangle|v_n(x_1)\rangle \tag{4.2}$$

where $|\varphi(x_2)\rangle$ corresponds to system II.

EPR then argued that a measurement on system I will not disturb system II as the two systems no longer interact (this is their implicit locality assumption). First consider a measurement of A on system I. This will not disturb system II. If the result of this measurement is a_k then the state of the system is reduced to

$$|\Psi(x_1,x_2)\rangle = |\psi_k(x_2)\rangle|u_k(x_1)\rangle \tag{4.3}$$

by this measurement. System I is now described by the state $|u_k(x_1)\rangle$ and system II is described by the state $|\psi_k(x_2)\rangle$. Now, instead consider a measurement of B on system I. This also will not disturb system II. If the result of this measurement is b_r then the state of the system is reduced to

$$|\Psi(x_1,x_2)\rangle = |\varphi_r(x_2)\rangle|v_r(x_1)\rangle \tag{4.4}$$

by this measurement. System I is now described by the state $|v_r(x_1)\rangle$ and system II is described by the state $|\varphi_r(x_2)\rangle$.

As system II has not been disturbed, no real change in system II can be said to have taken place. We find here, however, that this system can be described by two state vectors $|\psi_k(x_2)\rangle$ and $|\varphi_r(x_2)\rangle$ which are different. However, $|\psi_k(x_2)\rangle$ and $|\varphi_r(x_2)\rangle$ may be eigenstates of two different observables Q and P respectively which do not commute with corresponding eigenvalues p_k and q_r. We can predict with certainty that the result of measuring P will be p_k and that the result of measuring Q will be q_r. Therefore, using the criterion for the existence of elements of reality we see that there exist elements of reality corresponding to both P and Q. However, we saw above that if two physical properties corresponding to two non-commuting operators can be shown to have simultaneous physical reality then quantum mechanics must be incomplete. EPR concluded their paper by stating their belief that a more complete description of physical reality that quantum mechanics is possible in a more advanced theory. The completeness of quantum mechanics was defended by Bohr (1935). He states that, by choosing to measure one property of system I, there is "an influence on the very conditions which define the possible types of predictions regarding the future behaviour of the system." From a realist point of view this amounts to a denial of locality but it is not clear whether or not Bohr believed in nonlocality in quantum mechanics.

EPR used a particular entangled state which contained a position and a momentum correlation between the two particles. Bohm considered a simpler state, the singlet state, which involves spin correlation between the two spin $\frac{1}{2}$ particles when he presented the argument in 1951. The singlet state can be written

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\Big(|+\rangle_{1x}|-\rangle_{2x} - |-\rangle_{1x}|+\rangle_{2x}\Big) \tag{4.5}$$

where $|\pm\rangle_{ix}$ is the state of particle i when it has spin $\pm\frac{1}{2}$ along the x axis. This can be written in terms of y spin components instead:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\Big(|+\rangle_{1y}|-\rangle_{2y} - |-\rangle_{1y}|+\rangle_{2y}\Big) \tag{4.6}$$

We can now run the EPR argument using this state: If the two particles interact such that their state becomes a singlet state and then separate but remain in a singlet state then one can consider measurements of spin along the x axis and along the y axis. If one measures the x component of spin of particle 1 then from equation (4.5) the particle 2 is predicted to have the opposite value for spin along the x axis. Likewise with the y spin measurements. Therefore, both the x and y components of spin of particle 2 have simultaneous physical reality and so the quantum mechanical description is incomplete.

4.3 Bell's theorem

As we have already noted the EPR argument makes a locality assumption. In this section we will review Bell's theorem which demonstrates that this assumption cannot always be true in quantum theory. Bell's theorem applies to what have come to be called local realistic models or local hidden-variable models. Locality is the assumption that a measurement on one system cannot affect another distant system. Realism in the context of Bell's theorem is the assumption that there exist hidden variables that, at least partially, determine the outcome of a measurement. For the moment we will assume that

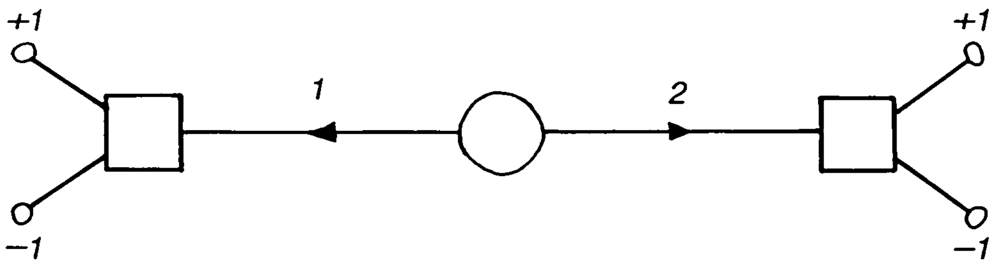

Fig. 4.1. Apparatus used to run Bell's demonstration of nonlocality.

the hidden variables exactly determine the outcomes of measurements, that is we will assume determinism. We will discuss later indeterministic hidden-variable interpretations We take two systems, 1 and 2 which interact for some time and then separate to impinge on measurement apparatuses at space-like separated locations (see fig. 4.1). The state of the total system 1 and 2 is given in quantum mechanics by some state vector $|\psi\rangle$. However, this is not necessarily a complete description of the state of the two particles. Thus we suppose that a complete description of the two particles is given by set of hidden variables λ. (This nomenclature is slightly misleading as the hidden variables offer a complete description and therefore may contain information that is not 'hidden' such as the quantum mechanical state vector.) As expressed by Bell: "It is a matter of indifference . . . whether λ denotes a single variable or a set or even a set of functions, and whether the variables are discrete or continuous." We make a measurement of $A(a)$ on system 1 and a measurement of $B(b)$ on system 2. a and b are local variables, that is they can be set locally in the region of the measurement of A and B respectively. In the example considered by Bell the measurements are measurements of spin and the local variable is the direction along which the spin is measured. We will assume that the measurements A and B each have two possible outcomes, either $+1$ or -1.

The assumption of locality demands that the result of measuring A on particle 1 does not depend on the setting of the local variable b of particle 2. Likewise, the result of measuring B on particle 2 does not depend on the setting of the local variable a of particle 1. Hence the result of a measurement of A is determined by a and λ and the result of measuring B is determined by b and λ such that we have the result functions

$$A(a, \lambda) = \pm 1 \tag{4.7}$$

$$B(b, \lambda) = \pm 1 \tag{4.8}$$

For an ensemble of systems the expectation value of $A(a)B(b)$ is

$$E(a,b) = \int A(a,\lambda)B(b,\lambda)\rho(\lambda)d\lambda \tag{4.9}$$

where $\rho(\lambda)$ is the probability distribution of λ over the ensemble. This expectation value should equal the quantum mechanical expectation value calculated for the state of the systems of the ensemble. However, we will see that this is not possible. The inequalities derived by Bell in 1964 made use of the special properties of a singlet state and assumed ideal conditions not realizable in a real experiment. Therefore, we will consider the inequalities derived by Clauser, Horne, Shimony, and Holt (CHSH) in 1969 which do not assume ideal conditions and are not specific to the singlet state. To derive these inequalities first we write,

$$E(a,b) - E(a,b') = \int \Big[A(a,\lambda)B(b,\lambda) - A(a,\lambda)B(b',\lambda)\Big]\rho(\lambda)d\lambda$$

Rearranging gives:

$$E(a,b) - E(a,b) = \int A(a,\lambda)B(b,\lambda)\Big[1 \pm A(a',\lambda)B(b',\lambda)\Big]\rho(\lambda)d\lambda \\ - \int A(a,\lambda)B(b',\lambda)\Big[1 \pm A(a',\lambda)B(b,\lambda)\Big]\rho(\lambda)d\lambda$$

Using (4.7) and (4.8), we get

$$|E(a,b) - E(a,b)| \leq \int \Big[1 \pm A(a',\lambda)B(b',\lambda)\Big]\rho(\lambda)d\lambda - \int \Big[1 \pm A(a',\lambda)B(b,\lambda)\Big]\rho(\lambda)d\lambda \tag{4.10}$$

Using (4.9) we obtain

$$|E(a,b) - E(a,b')| \leq \pm[E(a',b') + E(a',b)] + 2 \tag{4.11}$$

This gives the CHSH Bell inequalities

$$-2 \leq S(a,b,a',b') \leq 2 \tag{4.12}$$

where

$$S(a,b,a',b') = E(a,b) - E(a,b') + E(a',b) + E(a',b') \tag{4.13}$$

These are the inequalities most commonly used in discussions of Bell's theorem.

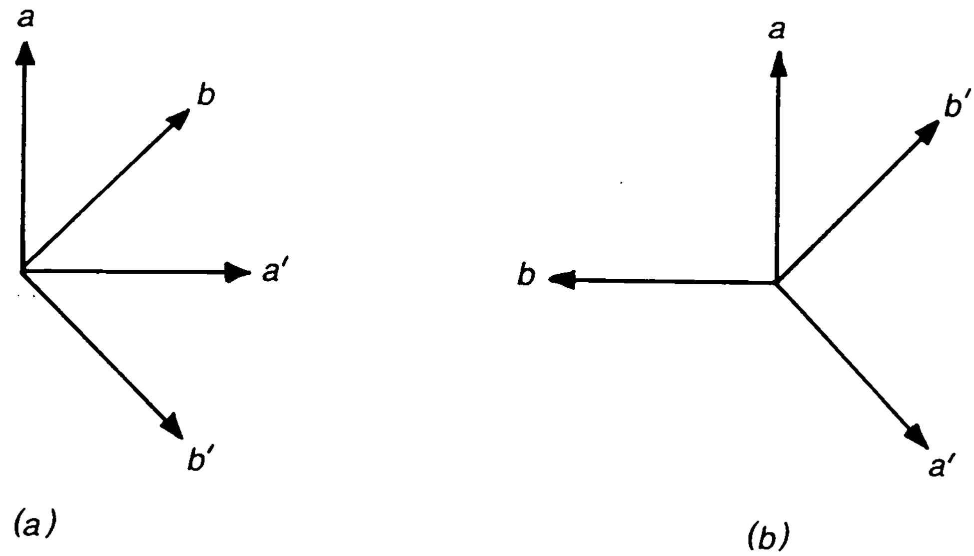

Fig. 4.2. Settings for maximum violation of (a) the upper limit and of (b) the lower limit of the CHSH inequalities for the singlet state.

Now consider a pair of spin $\frac{1}{2}$ particles prepared in the singlet state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\Big(|+\rangle_1|-\rangle_2 - |-\rangle_1|+\rangle_2\Big) \tag{4.14}$$

Suppose that these two particles separate, particle 1 moving along the z axis in the +ve direction and particle 2 moving along the same axis in the −ve direction. Then suppose that measurements of spin are made, $A(a)$ on particle 1 and $B(b)$ on particle 2, along directions at angles a and b respectively to the x axis in the xy plane. The possible results of these measurements are ±1 in appropriate units. The expectation value of $A(a)B(a)$ according to quantum mechanics is given by

$$E(a, b) = \langle\Psi|\sigma_1(a)\sigma_2(b)|\Psi\rangle \tag{4.15}$$

Using elementary quantum mechanics this gives

$$E(a, b) = -cos(a - b) \tag{4.16}$$

The maximum value of S is obtained when we put $a = 0$, $b = \frac{\pi}{4}$, $a' = \frac{\pi}{2}$, and $b' = \frac{3\pi}{4}$ (fig. 4.2(a)):

$$S(0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}) = 2\sqrt{2} \tag{4.17}$$

which violates the upper limit of the CHSH inequality. The minimum value of S is obtained when we put $a = 0$, $b = \frac{3\pi}{4}$, $a' = \frac{3\pi}{2}$, and $b' = \frac{\pi}{4}$ (fig 4.2(b)):

$$S(0, \frac{3\pi}{4}, \frac{3\pi}{2}, \frac{\pi}{4}) = -2\sqrt{2} \tag{4.18}$$

which violates the lower limit of the CHSH inequality. Therefore, we conclude that quantum mechanics is not a local realistic theory. It has been shown by Gisin (1991) that any pure entangled state will violate the CHSH inequalities although mixtures of entangled states do not necessarily violate these inequalities (see Popescu and Rohrlich (1992) and Braunstein, Mann and Revzen (1992))

4.4 A new way to obtain Bell inequalities

Whilst the above method of obtaining Bell inequalities is sufficient to demonstrate the contradiction between local realism and quantum mechanics, any intuitive understanding of the contradiction is lost in the derivation of the inequalities. In this section we will obtain the inequalities in a different way such that we do not lose sight of the origin of the contradiction (see also Hardy (1991d)). This approach is motivated by the demonstration of Bell's theorem due to Greenberger, Horne, and Zeilinger (1989), simplified by Mermin (1990b). In their approach in which more than two particles are used it is possible to demonstrate Bell's theorem by means of a direct contradiction without using any inequalities. We will discuss the GHZ argument in more detail in chapter 7. Following Braunstein and Caves (1989) and (1990) we will consider more than two settings of the local variables at each end of the apparatus.

Consider the following statements,

$$\begin{array}{ll}
\text{s1.} & A(a_1,\lambda)B(b_2,\lambda) = -1 \\
\text{s2.} & A(a_3,\lambda)B(b_2,\lambda) = -1 \\
\text{s3.} & A(a_3,\lambda)B(b_4,\lambda) = -1 \\
\text{s4.} & A(a_5,\lambda)B(b_4,\lambda) = -1 \\
\vdots & \vdots \\
\text{s}K-1. & A(a_{K-1},\lambda)B(b_K,\lambda) = -1 \\
\text{s}K. & A(a_1,\lambda)B(b_K,\lambda) = +1
\end{array}$$

where K is even. Each quantity $A(a_k, \lambda)$ and $B(b_k, \lambda)$ appears twice on the LHS.

Hence, the product of all these equations must be equal to $+1$ on the LHS but equal to -1 on the RHS. Therefore, if all the statements s1 to sK are true then there is a contradiction. If we choose to measure $A(a_k)$ then we cannot also measure $A(a_l)$ for the same value of λ if $a_l \neq a_k$. Consequently, we cannot actually measure all of the quantities in the statements s1 to sK. However, by considering the probabilities associated with an ensemble of experiments, we will be able to show that all of the statements s1 to sK must be true for some of the experiments (i.e. for some values of λ) when certain conditions are met. Let the probability $p^{\pm}(a,b)$ be equal to the number of experiments for which $A(a)B(b) = \pm 1$ divided by the total number of experiments in the limit as the total number of experiments tends to infinity. This probability is given by the predictions of quantum theory. The probabilities for each of the statements s1 to sK being true are

s1	$p_1^- = p^-(a_1, b_2)$
s2	$p_2^- = p^-(a_3, b_2)$
s3	$p_3^- = p^-(a_3, b_4)$
s4	$p_4^- = p^-(a_5, b_4)$
.	.
.	.
.	.
s$K-1$	$p_{K-1}^- = p^-(a_{K-1}, b_2)$
sK	$p_K^+ = p^+(a_1, b_K)$

These probabilities are not independent and consequently we cannot write down an equation for the probability that all the statements s1 to sK are true. However we can write down an inequality. If P is the probability that all the statements, s1 to sK are true then $1 - P$ is the probability that one or more of the statements is false. The probability that one or more of the statements is false must be less than or equal to the sum of the probabilities for each individual statement being false. The previous sentence represents a crucial step in the argument and so should be understood. It can be expressed mathematically by

$$1 - P \leq \sum_{k=1}^{K-1}(1 - p_k^-) + (1 - p_K^+) \tag{4.19}$$

This simplifies to

$$P \geq \sum_{k=1}^{K-1} p_k^- + p_K^+ - (K-1) \tag{4.20}$$

P is a statistical probability (the number of runs of the experiment for which s1 to sK are true divided by the total number of runs of the experiment). Therefore, if $P > 0$ then there will certainly be some runs of the experiment for which all the statements s1 to sK are true leading to the previously mentioned contradiction. From inequality (4.20) we see that $P > 0$ if

$$\sum_{k=1}^{K-1} p_k^- + p_K^+ > K - 1 \tag{4.21}$$

If this inequality is satisfied then there is a contradiction between quantum mechanics and local realism. Usually, in Bell theory, we express the contradiction the other way round, thus reversing the above inequality gives the Bell inequality

$$\sum_{k=1}^{K-1} p_k^- + p_K^+ \leq K - 1 \tag{4.22}$$

If this inequality is violated by quantum mechanics then there is a contradiction between quantum mechanics and local realism. We will now show (i) that inequality (4.22) can be violated by quantum mechanics and, (ii) that there is a local theory which saturates the inequality demonstrating that it is a good inequality for expressing the contradiction between quantum mechanics and local realism.

(i) For the singlet state considered in the previous section we have, according to quantum mechanics

$$p^{\pm}(a,b) = \frac{1}{2}\Big(1 \mp \cos(a-b)\Big) \tag{4.23}$$

This can easily be calculated from equation (4.16) using

$$E(a,b) = p^+(a,b) - p^-(a,b) \tag{4.24}$$

and

$$p^+(a,b) + p^-(a,b) = 1 \tag{4.25}$$

We can choose the angles a_1, b_2, a_3,. . . b_K to be evenly spread (see fig. 4.3) so that

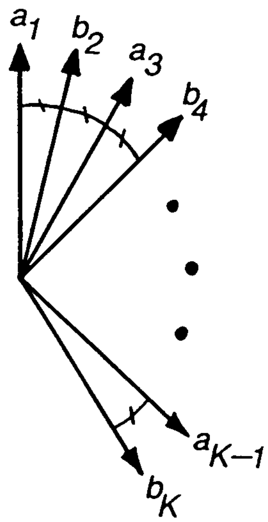

Fig. 4.3. Settings a_1 to b_K evenly spread.

$$b_K - a_1 = \phi \tag{4.26}$$

$$b_n - a_{n\pm 1} = \mp \frac{\phi}{K-1} \tag{4.27}$$

Putting (4.23) into inequality (4.22) we get a maximum violation of the inequality when

$$\phi = \frac{(K-1)\pi}{K} \tag{4.28}$$

For $K = 4$ the inequality can be written

$$p_1^- + p_2^- + p_3^- + p_4^+ \leq 3 \tag{4.29}$$

Using equation (4.23) and putting $\phi = \frac{3\pi}{4}$ (from (4.28)) the LHS of (4.29) is equal to $2 + \sqrt{2}$ which violates (4.29). As $K \to \infty$ the LHS of inequality (4.22) tends to K. In this same limit, the RHS of inequality (4.20) tends to 1 and therefore $P \to 1$. This limit is particularly interesting. In the limit the inequality (4.22) is replaced by an the equality $P = 1$. Thus, in this sense, we get an inequality free demonstration of Bell's theorem in the limit of an infinite number of settings of the local variables. There is some similarity between this and the quantum Zeno effect (Misra and Sudarshan (1977)). If ϕ is small then from (4.23) we have

$$p^- = 1 - \frac{1}{2}\phi^2 \tag{4.30}$$

This means that $A(0, \lambda) = -B(\phi, \lambda)$ with probability equal to one to second order in ϕ. Similarly, we have $-B(\phi, \lambda) = A(2\phi, \lambda)$ with probability equal to one to second order in

ϕ. We can continue this from angle 0 to angle π establishing that $A(0,\lambda) = -B(\pi,\lambda)$ with probability equal to one in the limit as $\phi \to 0$, i.e. as $K \to \infty$. However, if $a = 0$ and $b = \pi$ then equation (4.23) tells us that $A(0,\lambda) = B(\pi,\lambda)$ with probability equal to one. Thus we have a contradiction. To see that this argument is rigorous we have to appeal to the mathematics above leading us to the inequality (4.22) as this is the correct mathematics to run such a limiting argument. In this limit the fact that we have $1 - \frac{1}{2}\phi^2$ rather than $1 - \frac{1}{2}\phi$ is crucial in obtaining the contradiction as we shall now see.

(ii) A local model for the spin experiment described above is given by Bell (1964) (see also page 86 of Redhead (1987)). Simplified, for the present purpose, this model assumes that the hidden variable λ is an angle in the interval $0 \leq \lambda \leq 2\pi$ with a uniform distribution over all the possible directions, ie $\rho(\lambda) = \frac{1}{2\pi}$. We put

$$A(a,\lambda) = \mathrm{sign}(a - \lambda) \tag{4.31}$$

$$B(b,\lambda) = \mathrm{sign}(\lambda - b) \tag{4.32}$$

The assumption of locality is satisfied by these choices. Using (4.31) and (4.32) and the fact that λ is evenly distributed we obtain

$$p^{+}(a,b) = \frac{|a-b|}{\pi} \tag{4.33}$$

$$p^{-}(a,b) = 1 - \frac{|a-b|}{\pi} \tag{4.34}$$

Making the same choices for $a_1, b_2, \ldots b_K$ as in equations (4.26) and (4.27) we find that the probabilities (4.33) and (4.34) saturate inequality (4.22). We see that this time $a - b = \phi$ is not squared in the expression for p^-. It is for this reason that we can avoid the contradiction discussed in (i) above.

The CHSH inequality can be obtained from inequality (4.22) by using (4.24) and (4.25) and putting $K = 4$. Hence, it is not surprising that quantum mechanics violates (4.22). However, the derivation of (4.22) is simpler and more intuitive than previous derivations of Bell inequalities. The contradiction between quantum mechanics and local realism can be understood in the following way: Quantum mechanics requires that a certain set of statements, s1 to sN, must all be true for some runs of the experiment even though, from the local realist point of view, they are clearly contradictory when they are all true.

4.5 The question of indeterminism

Until now we have only considered deterministic hidden-variable theories. We have seen that if these theories are also local then they cannot agree with quantum mechanics. Naturally, the question arises of whether it is the assumption of determinism or the assumption of locality that produces the contradiction with quantum mechanics. In this section we will first see that local hidden-variable theories must in fact be deterministic if they are to reproduce the results of quantum theory for parallel measurements of spin in an ideal experiment. Second, we will see that even if we are not in an ideal situation, then the following results holds: We can always reproduce the predictions of any local indeterministic hidden-variable theory by a local deterministic hidden-variable theory. Consequently, none of the class of local indeterministic hidden-variables theories can reproduce quantum mechanics for if they could then we could find a local deterministic hidden-variable theory that could reproduce quantum mechanics contradicting our above result. Thus, it is clearly the assumption of locality that produces the contradiction with quantum mechanics.

The first result is quite easy to see. By using EPR's reasoning, we can establish that result functions $A(a, \lambda)$ and $B(b, \lambda)$ must exist. The result of measuring $A(a)$ does not depend on the value of b if we assume locality. If a is set equal to b then from equation (4.16) we see that, in the case of the singlet state, that the result of measuring $A(a)$ must be equal to -1 times the result of measuring $B(b)$. Thus in this case there is no room for indeterminism. That is we must have

$$A(a, \lambda) = -B(a, \lambda) \tag{4.35}$$

where A is exactly determined. However, as the result of measuring A does not depend on the value of b, A must always be exactly determined even when $b \neq a$. Similar arguments apply to B. Thus locality requires determinism in order to satisfy the predictions of quantum mechanics in the case of measurements of spin along parallel directions. As stated by Bell (1964), "since we can predict in advance the result of measuring any chosen component of σ_2, by previously measuring the same component of σ_1, it follows that the result any such measurement must actually be predetermined." This shows that it is the assumption of locality that is the essential assumption in the demonstration of Bell's theorem. To run this argument to demonstrate determinism, we have had to assume ideal conditions, that is ideal measurement apparatus and an ideal preparation of the singlet state. To demonstrate that nature is really nonlocal we must show that the predictions of quantum mechanics are correct by performing an experiment. However, in a real experiment we cannot realize ideal conditions and therefore the above argument for determinism becomes redundant.

Our second approach which is originally due to Stapp (1980) and Fine (1982) remains valid in nonideal experiments. Instead of showing that local hidden-variable theories must be deterministic in order to have any chance of reproducing the predictions of quantum theory, we will allow consideration of indeterministic theories. Consider an indeterministic theory which gives probabilities for various outcomes. Thus, let $p_1^{\pm}(a,\lambda)$ be the probability that, for hidden variables λ, the result of measuring $A(a)$ is ± 1 and similarly let $p_2^{\pm}(b,\lambda)$ be the probability that the outcome of measuring $B(b)$ with hidden variables λ is ± 1. The assumption of locality has been incorporated into these probabilities because we have assumed that the probability of getting $A(a) = \pm 1$ does not depend on the setting of b and similarly that the probability of getting $B(b) = \pm 1$ does not depend on the setting of a. Furthermore, we assume that the stochastic process at each end is independent of the stochastic process at the other end. Thus, we can write

$$p_{12}^{\pm\pm}(a,b,\lambda) = p_1^{\pm}(a,\lambda)p_2^{\pm}(b,\lambda) \tag{4.36}$$

This is the condition written down by Clauser and Horne (1974) defining what they called objective local theories but which we will continue to call local realistic or local hidden-variable theories. We will now show that the predictions of such theories can always be reproduced by local deterministic hidden-variable theories. Consider a hidden variable theory in which, in addition to λ, there also exist hidden variables μ ($0 \leq \mu \leq 1$) with distribution $\rho_1(\mu) = 1$ pertaining to particle 1 and ν ($0 \leq \nu \leq 1$) with distribution $\rho_2(\nu) = 1$ pertaining to particle 2. The two distribution functions are taken to be independent of each other and independent of the distribution of λ, that is

$$\rho_{12}(\lambda,\mu,\nu) = \rho(\lambda)\rho_1(\mu)\rho_2(\nu) \tag{4.37}$$

where $\rho_{12}(\lambda,\mu,\nu)$ is the distribution function for λ, μ, and ν. Instead of the stochastic processes at each end being truly random we will suppose that these additional hidden variables serve to exactly determine the outcome at each end by the functions

$$p_1^{+}(a,\lambda,\mu) = \begin{cases} 1 & if \quad \mu \leq p_1^{+}(a,\lambda) \\ 0 & if \quad \mu > p_1^{+}(a,\lambda) \end{cases} \tag{4.38}$$

$$p_1^{-}(a,\lambda,\mu) = \begin{cases} 1 & if \quad 1-\mu \leq p_1^{-}(a,\lambda) \\ 0 & if \quad 1-\mu > p_1^{-}(a,\lambda) \end{cases} \tag{4.39}$$

$$p_2^+(b,\lambda,\nu) = \begin{Bmatrix} 1 & if & \nu \leq p_2^+(b,\lambda) \\ 0 & if & \nu > p_2^+(b,\lambda) \end{Bmatrix} \tag{4.40}$$

$$p_2^-(b,\lambda,\nu) = \begin{Bmatrix} 1 & if & 1-\nu \leq p_2^-(b,\lambda) \\ 0 & if & 1-\nu > p_2^-(b,\lambda) \end{Bmatrix} \tag{4.41}$$

These probabilities are always either 0 or 1 confirming that we have constructed a deterministic hidden-variable theory. Also, p_1 does not depend on ν or b and p_2 does not depend on μ or a demonstrating that it is in fact a local deterministic hidden-variable theory. It is clear that

$$p_1^\pm(a,\lambda) = \int p_1^\pm(a,\lambda,\mu)\rho_1(\mu)d\mu \tag{4.42}$$

and

$$p_1^\pm(a,\lambda) = \int p_1^\pm(a,\lambda,\mu)\rho_1(\mu)d\mu \tag{4.43}$$

and furthermore, as the distributions of μ and ν are independent of each other and of the distribution of λ we can satisfy the factorisability condition (4.36). Therefore, we have succeeded in producing a local deterministic-hidden variable which can reproduce the predictions of any local indeterministic hidden-variable theory and consequently, for the reasons stated above, we see that the predictions of quantum mechanics cannot be reproduced by any local hidden-variable theory, deterministic or not.

This discussion motivates the following approach to Bell's theorem. First we suppose that a complete description of a system is in fact provided by its state vector $|\Psi\rangle$. This requires that we identify the state vector with the hidden-variables because we have supposed that these give a complete description of the system. However, for an entangled state such as the singlet state we notice immediately that

$$p_{12}^{\pm\pm}(a,b,\lambda=|\Psi\rangle) = \langle\Psi|\widehat{P}_1^\pm(a)\widehat{P}_2^\pm(b)|\Psi\rangle \neq p_1^\pm(a,\lambda=|\Psi\rangle)p_2^\pm(b,\lambda=|\Psi\rangle) \tag{4.44}$$

where $\widehat{P}_1^\pm(a)$ and $\widehat{P}_2^\pm(b)$ are projection operators onto the states with $A(a) = \pm 1$ and $B(b) = \pm 1$ respectively. This tells us that a local hidden-variable theory in which $|\Psi\rangle$ is the hidden variable is not possible because (4.44) contradicts (4.36). The task is then to see if by supplementing the statevector with additional hidden variables (or replacing it altogether with some other hidden-variable description) we can restore locality such that condition (4.36) is satisfied. We see that determinism is the best chance we have because, from the first argument above, it is only with determinism can we reproduce the predictions of quantum theory for parallel spin measurements in the ideal singlet case and from the second argument we see that anything that indeterminism can do determinism can also do. However, even complete determinism fails when we attempt to reproduce the whole range of predictions for spin measurements on the singlet state. Given these arguments it is natural to ask why we have stated that Bell's theorem only applies to local realistic theories, why not just to any local theory? The necessity of the assumption of realism can be understood because it is necessary that we can admit into our analysis the concept of what would have happened if some other measurement measurement had been performed (that is, other than the actual measurement that was performed). Inspection of the derivation of the Bell inequalities makes it clear that we have assumed that these 'what if' values have the same distribution as the measured values. The connection with the reality criterion of EPR should be made clear: this also assumes that, under certain circumstances, there exists an element of reality even when there is no measurement of the corresponding physical quantity.

Chapter 5

Experiments to Test Local Realism

5.1 Introduction

The singlet state was used in the previous chapter as an example to illustrate quantum nonlocality. In fact there are many different situations in which quantum nonlocality can be manifest. Discussion of new experiments to test local realism can serve two purposes. First, realizable experimental proposals offer the experimentalists the chance to demonstrate quantum nonlocality in new situations with the possibility of closing some of the loop holes open to local realist theories (Zeilinger (1986), Pascazio and Reignier (1987), Squires (1990), Santos (1991) and Lepore and Selleri (1990)). Secondly, even if the experiments are only possible at the gedanken level they may serve to further clarify the phenomenon of nonlocality.

With present technology it is not possible to test local realism without making some supplementary assumptions. These supplementary assumptions would not be necessary if the particles, when leaving the source, had well correlated directions and if the detector efficiencies were sufficiently high. In the next section we will discuss the application of Bell inequalities to real experiments and the way in which the supplementary assumptions are incorporated into the inequalities. Then we will discuss the problem of directional correlation of the photons produced in atomic cascades of the kind used in many experiments that have been performed to test Bell's inequalities. We will see that better results can, in principle, be obtained by using a new source that employs two nonlinear crystals to produce photon pairs by parametric down conversion. A way of obtaining the singlet state by 'shuffling boxes' will also be discussed. This method will be seen to be useful in preparing other entangled states in later chapters although only at the gedanken level.

5.2 Real experiments and supplementary assumptions

In the previous chapter we considered an experiment in which $A(a)$ ($B(b)$) is measured on particle 1 (2) with possible outcomes $+1$ or -1. This assumes that the apparatuses registers a result for every particle pair. In a real experiment this will not be the case and so we have the additional possibility that the outcome is 0 denoting that the apparatus has failed to register a result. However, in the derivation of the CHSH Bell inequalities in the previous chapter, we did not take the possibilities $A(a,\lambda) = 0$ and $B(b,\lambda) = 0$ into account. In fact, examination of the derivation of these inequalities, particularly the step leading to (4.10), reveals that these inequalities still hold when we include the possibility of null results. Thus the CHSH inequalities remain valid when applied to real experiments. Unfortunately, it has not yet been possible to realize the experimental conditions under which quantum mechanics predicts a violation of these inequalities. This is because the proportion of runs of the experiment in which both particles are detected is small for two reasons: (i) The directions in which the particles of a pair are emitted from the source are not well correlated and consequently not all of the particles fall on the appropriate detectors. (ii) The quantum efficiencies of the detectors are low such that even when the particles do fall on the appropriate detectors they may not actually be detected. As we shall see, the technology to overcome the first of these problems already exists and in fact an experiment has been performed in which directional correlation was not a problem. However, to date no experiments have been performed which have overcome the second problem and consequently the CHSH inequalities have not been tested in a real experiment. If detectors with efficiencies greater than 83% were to become available then a direct test of the CHSH inequalities would be possible (Ou (1988)). The same comments apply to other Bell inequalities that have been derived (that do not employ any supplementary assumptions), for example, the Clauser and Horne (1974) inequalities.

To overcome these problems in the experiments that have been performed it has been necessary to make a supplementary assumption and use this in deriving a new set of Bell-type inequalities. Such a supplementary assumption was first introduced by Clauser, Horne, Shimomy, and Holt (1969). See also Clauser and Horne (1974). First we will derive a set of inequalities and then introduce the supplementary assumption.

If

$$
\begin{aligned}
-X &\leq x \leq X \\
-X &\leq x' \leq X \\
-Y &\leq y \leq Y \\
-Y &\leq y' \leq Y
\end{aligned}
\tag{5.1}
$$

where, necessarily

$$
\begin{aligned}
X &\geq 0 \\
Y &\geq 0
\end{aligned}
\tag{5.2}
$$

Then it can be shown that

$$-2XY \leq xy + x'y + xy' - x'y' \leq 2XY \tag{5.3}$$

To see this first consider the case $y \geq y'$. For this case it follows that, using (5.1)

$$
\begin{aligned}
&x(y+y') + x'(y-y') \leq x(y+y') + X(y-y') \\
=&y(x+X) + y'(x-X) \leq Y(x+X) - Y(x-X) = 2XY
\end{aligned}
$$

If instead we have $y' \geq y$ then

$$
\begin{aligned}
&x(y+y') + x'(y-y') \leq x(y+y') - X(y-y') \\
=&y(x-X) + y'(x+X) \leq Y(x-X) - Y(x-X) = 2XY
\end{aligned}
$$

Thus the upper limit of (5.3) is established. That the lower limit is also satisfied is easily seen by multiplying (5.3) by -1 and making the transformations $x \rightarrow -x$ and $x' \rightarrow -x'$ which has the effect of transforming the previous lower limit into the same form as the upper limit. However, the transformation of x and x' leaves the inequalities (5.1) unaffected and consequently the lower limit of (5.3) is must be satisfied as the upper limit is satisfied.

Now let x, x', and X be some quantities measured at end 1. It does not matter what these quantities are as long as they satisfy inequalities (5.1) and (5.2). If the hidden variables are λ then the actual values of these quantities are determined by some functions $x(\lambda)$, $x'(\lambda)$, and $X(\lambda)$. Similarly, let y, y', and Y be quantities that can be measured at end 2 with values given by functions $y(\lambda)$, $y'(\lambda)$, and $Y(\lambda)$. This incorporates the assumption of locality because the value of a quantity at one end does not depend on the choice of measurement at the other end. The expectation values of these quantities are

$$\langle xy\rangle = \int x(\lambda)y(\lambda)\rho(\lambda)d\lambda \tag{5.4}$$

etc. By taking the expectation value of inequality (5.3) and then dividing by $\langle XY\rangle$ we get

$$-2 \leq \frac{\langle xy\rangle}{\langle XY\rangle} + \frac{\langle x'y\rangle}{\langle XY\rangle} + \frac{\langle xy'\rangle}{\langle XY\rangle} - \frac{\langle x'y'\rangle}{\langle XY\rangle} \leq 2 \tag{5.5}$$

At this stage we have not made any supplementary assumptions (that is, in addition to locality). As long as the inequalities (5.1) and (5.2) are satisfied by the quantities we choose then inequalities (5.4) can be regarded as true Bell inequalities. For example, if we put $x = A(a)$, $x' = A(a')$, $y = B(b)$, $y' = B(b')$, and $X = Y = 1$ then the inequalities (5.1) and (5.2) are satisfied and inequalities (5.5) are then the CHSH inequalities considered in the previous chapter. In the next chapter we will also make use of inequalities (5.5) in a way that does involve supplementary assumptions and in two ways that do not involve supplementary assumptions. For our present purposes we will see how supplementary assumptions can be incorporated into the analysis of the type of experiments discussed so far and to be discussed in the rest of this chapter in order to make experiments possible with existing technology. As we are now taking the measurement apparatuses to be nonideal we will assume that in addition to the hidden variables λ which give a complete description of the quantum particles, there also exist hidden variables μ and ν associated with the measurement apparatuses at end 1 and end 2 respectively. This is similar to section 4.5. As in section 4.5 we will assume that the distributions $\rho(\lambda)$, $\rho_1(\mu)$, $\rho_2(\nu)$ of λ, μ, and ν are independent, that is

$$\rho_{12}(\lambda, \mu, \nu) = \rho(\lambda)\rho_1(\mu)\rho_2(\nu) \tag{5.6}$$

where $\rho_{12}(\lambda, \mu, \nu)$ is the distribution function for λ, μ, and ν. This assumption is justified as the quantum particles and the measurement apparatuses are initially noninteracting. The result of measuring $A(a)$ is given by the function $A(a, \lambda, \mu)$ and the result of measuring $B(b)$ is given by the function $B(b, \lambda, \nu)$. Let $p_1^{\pm}(a, \lambda)$ be the probability of getting $A(a) = 1$ given that the hidden variables of the quantum particles are λ. That is

$$p_1^{\pm} = \int \frac{1}{2}\big(|A(a, \lambda, \mu)| \mp A(a, \lambda, \mu)\big)\rho_1 d\mu \tag{5.7}$$

Similarly, for particle 2 we have

$$p_2^{\pm} = \int \frac{1}{2}\big(|B(b, \lambda, \nu)| \mp B(b, \lambda, \nu)\big)\rho_2 d\nu \tag{5.8}$$

The supplementary assumption can now be written

$$p_1^+(a,\lambda) + p_1^-(a,\lambda) = p_1(\lambda)$$
$$p_2^+(b,\lambda) + p_2^-(b,\lambda) = p_2(\lambda) \tag{5.9}$$

The point to notice about these equations is that their RHS is independent of the setting of the local variable a or b. The assumption can be stated in words in the following way: The total probability of detecting a particle at each end is independent of the setting of the local variable. We can now set

$$x(\lambda) = p_1^+(a,\lambda) - p_1^-(a,\lambda) = A_1(a,\lambda) \tag{5.10}$$

$$x'(\lambda) = p_1^+(a',\lambda) - p_1^-(a',\lambda) = A_1(a',\lambda) \tag{5.11}$$

$$y(\lambda) = p_2^+(b,\lambda) - p_2^-(b,\lambda) = B_2(b,\lambda) \tag{5.12}$$

$$y'(\lambda) = p_2^+(b,\lambda) - p_2^-(b,\lambda) = A_2(b',\lambda) \tag{5.13}$$

and

$$X(\lambda) = p_1(\lambda) \tag{5.14}$$

$$Y(\lambda) = p_2(\lambda) \tag{5.15}$$

These choices satisfy inequalities (5.1) and (5.2). The essential role played by the supplementary assumption (5.9) should be understood: With the choices (5.10) to (5.15), it is only possible for the same X to appear in each of the first two inequalities of (5.1) by virtue of the fact that $p_1(\lambda)$ is independent of a (or a'). Substituting (5.10) to (5.15) into (5.5) and using (5.6) to (5.8) we obtain

$$-2 \leq E(a,b) + E(a',b) + E(a,b') - E(a',b') \leq 2 \tag{5.16}$$

where the correlation function $E(a,b)$ is given by

$$E(a,b) = \frac{\langle A(a)B(b)\rangle}{\langle |A(a)B(b)|\rangle} \tag{5.17}$$

The denominator of (5.17) is equal to the probability of getting a non-null result at both ends.

Inequalities (5.16) with the correlation function (5.17) are essentially the same as the CHSH inequalities considered in the previous chapter except that only those runs of the experiment for which both particles are detected contribute to $E(a,b)$. The supplementary assumption essentially amounts to assuming that those runs of the experiment for which both particles are detected represent a fair sample of events. With this assumption the expectation value $E(a,b)$ that would be measured with ideal conditions can be taken to be equal to the expectation value formed by only considering those runs of the experiment in which both particles are detected and it is the latter values that are then substituted into the inequalities (5.16). It is not possible to test the supplementary assumption empirically except in the case of an ideal experiment because then the probability of getting a detection at end 1, say, is equal to 1 and this could be checked by noticing that a particle is detected for every run of the experiment. Similar comments apply to particle 2. Consequently, the RHS of (5.9) is then trivially independent of the setting of the local variables and thus, (5.9) need not be regarded as an assumption in the ideal case.

5.3 Photon sources with correlated polarisations

Most of the experiments performed to test local realism have employed photon pairs prepared in a state in which they have correlated polarisations. This is analogous to the case of two spin half particles prepared in the singlet state discussed in the previous chapter.

In experiments to test local realism we can use pairs of photons with correlated polarisations. (This is analogous to the case of two spin half particles prepared in the singlet state discussed in the previous chapter.) In most experiments conducted so far these pairs have been created in atomic cascades. For example the $\mathbf{J}=0 \rightarrow \mathbf{J}=1 \rightarrow \mathbf{J}=0$ decay was used in the series of experiments conducted at Orsay (Aspect, Grangier and Roger (1981) and (1982) and Aspect, Dalibard and Roger (1982)). Cascades have also been used by Freedman and Clauser (1972), Holt and Pipkin (1974), Clauser (1976), and Fry and Thompson (1976). This process is essentially a three body decay (two photons and the atom) and consequently the directions in which the two photons are emitted are poorly correlated. This means that, even with 100% efficient detectors we cannot use such a source to test Bell's inequalities without using the supplementary assumption. The reason for this is that in most events only one or none of the two photons will impinge on the appropriate detectors. The supplementary assumption makes it possible to consider only those events in which the both photons are detected by the appropriate detectors. Santos (1991) has actually constructed a local realistic model which reproduces the predictions

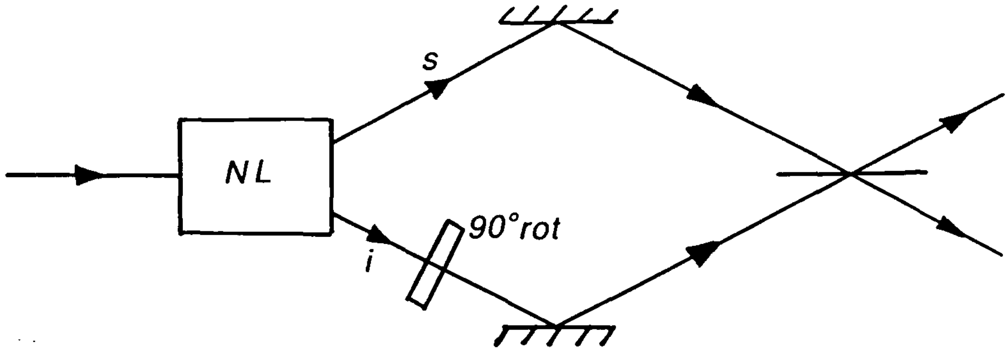

Fig. 5.1. Photon pairs are created by parametric down conversion. The polarisation of one photon is rotated by 90° and then the two modes are mixed at a beam splitter.

of quantum mechanics for cascade sources even with 100% efficient detectors (see also Santos (1990)). This model violates the supplementary assumption.

An alternative source of photon pairs with correlated polarisations uses degenerate parametric down conversion. This produces pairs of photons with the same polarisation and the same frequency. The polarisation of one of the photons is rotated through 90° and then the two modes are mixed at a beam splitter (see fig. 5.1). This experiment has been performed by Alley and Shih (1987) and by Ou and Mandel (1988). Unfortunately, in half of the events, both photons will go in the same direction at the beam splitter and in the other half of the events they will go in opposite directions. Hence, again we do not have sufficient correlation between the directions of the photons.

The problem of poor directional correlation can be overcome by performing a different type of experiment, namely, a two photon interference experiment using pairs of photons created in parametric down conversion. Horne Shimony and Zeilinger (1989) have proposed an experiment of this type which has been performed by Rarity and Tapster (1990). Other experiments not involving a singlet type state have been proposed by Żukowski and Pykacz (1988), Grangier, Potasek and Yurke (1988), Ou (1988), Żukowski (1990), and Żukowski and Zeilinger (1991). The two photons emitted in parametric down conversion have correlated directions (Morrow (1973) and Hong and Mandel (1985)) and consequently, it is possible to arrange the apparatus such that each photon always impinges on one of the appropriate detectors. If the quantum efficiencies of the detectors are greater than 83% then a test of Bell's inequalities without auxiliary assumptions is possible (Ou (1988)).

Here we will propose an approach which combines the old (correlated polarisations) and the new (correlated directions) (also see Hardy (1992a)). We propose a source that emits pairs of photons in the state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\Big(|\updownarrow\rangle_{\mathbf{k}_s}|\updownarrow\rangle_{\mathbf{k}_i} + |\leftrightarrow\rangle_{\mathbf{k}_s}|\leftrightarrow\rangle_{\mathbf{k}_i}\Big) \tag{5.18}$$

where $|\updownarrow\rangle_{\mathbf{k}_{s,i}}$ and $|\leftrightarrow\rangle_{\mathbf{k}_{s,i}}$ are the states of a photon with polarisation along the x and y axes respectively with wave vector $\mathbf{k}_{s,i}$. The z-axis for each photon is along the $\mathbf{k}_{s,i}$ direction. By measuring the polarisation of each photon at two distant systems placed in the paths of photons s and i it is possible, with such a source, to get a direct test of Bell's inequalities without auxiliary assumptions if the detectors have quantum efficiencies greater than 83%.

The proposed source is shown in fig. 5.2. A laser beam is split at a beam splitter to form two mutually coherent classical beams with complex amplitudes $V_1(t)$ and $V_2(t)$ which impinge on two similar nonlinear crystals NL1 and NL2 creating photon pairs by parametric down conversion at each crystal. All the photons created in this way will have the same polarisation along some direction which we will call the x direction (i.e. along the x axis). A photon pair created at nonlinear crystal j ($j = 1, 2$) consists of a signal photon s_j and an idler photon i_j. The signal photons from NL1 and NL2 have the same frequency and likewise, the idler photons from NL1 and NL2 have the same frequency. The photons from NL1 each pass through a 90° polarisation rotator so that they now have polarisation along the y axis and are then fed into NL2 such that s_1 is aligned with s_2 to form the s beam and i_1 is aligned with i_2 to form the i beam. This idea of aligning trajectories is due to Z. Y. Ou (see acknowledgements of Zou, Wang and Mandel (1988). The comlex amplitudes V_1 and V_2 are chosen such that the rates of creation of pairs at NL1 and NL2 are the same and also sufficiently small such that there is a very small probability that two or more photon pairs will be created that cannot be time resolved by the detectors. Consequently, if two photons are detected, one in the s beam and one in the i beam then both of these photons must have come either from NL1 in which case they both have y polarisation, or from NL2 in which case they both have x polarisation. If measurements of polarisation are made along some direction other than the x or y direction then we do not distinguish between photon pairs that have come from NL1 and photon pairs that have come from NL2. We will show that the state of such photon pairs can be that in (5.18).

The parametric interaction at the nonlinear crystal j is

$$H_{I_j} = \hbar g_j V_j(t)\widehat{s}^{\dagger}_{x_j}\widehat{i}^{\dagger}_{x_j} + \text{H.c.} \tag{5.19}$$

where $\widehat{s}_{x_j}$ and $\widehat{i}_{x_j}$ are the annihilation operators for photons s_j and i_j respectively with polarisation along the x axis (see Zou et al. (1991)). g_j is a frequency proportional to

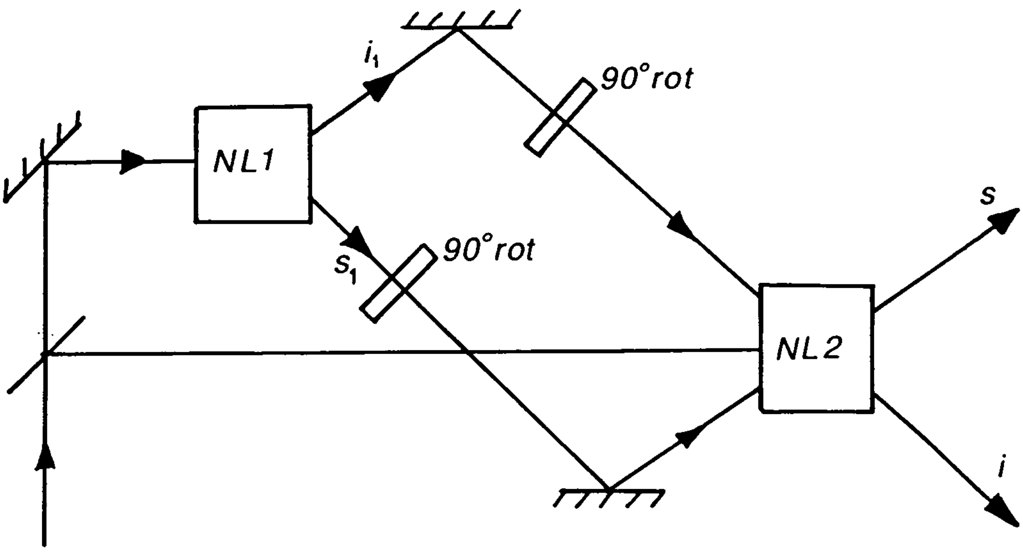

Fig. 5.2. Two nonlinear crystals arranged to produce pairs of photons with correlated polarisations.

the nonlinear susceptibility of the nonlinear crystal and the complex amplitude $V_j(t)$ is dimensionless. The first term on the RHS of (5.19) gives rise to the creation of the signal and idler photons. The second term gives rise to the annihilation of photons in modes s_{x_j} and i_{x_j} (see page 291 of Klyshko (1988)). It may be thought that this second term will lead to some of the photons from NL1 being annihilated in NL2. However, this will not happen because the photons from NL1 pass through 90° polarisation rotators so that they have annihilation operators $\widehat{s}_{y_j}$ and $\widehat{i}_{y_j}$ which do not appear in (5.19). The process of parametric down conversion creates pairs of photons with the same polarisation along some direction which we have called the x direction. Now, H_{I_j} must be Hermitian and therefore it will not contain terms which lead to the annihilation of photons s_1 and i_1 with polarisation along the y axis.

If the initial state is the vacuum then, using the interaction picture with the interaction given in equation (5.19) and neglecting terms with more than two photons, we get for the final state,

$$|\Psi\rangle = |vac\rangle + f_1 V_1(t) e^{i(\alpha_s+\alpha_i)} |0\rangle_{s_x} |1\rangle_{s_y} |0\rangle_{i_x} |1\rangle_{i_y} + f_2 V_2(t+\tau) |1\rangle_{s_x} |0\rangle_{s_y} |1\rangle_{i_x} |0\rangle_{i_y} \quad (5.20)$$

where $|f_j|^2$ is the fraction of the incident light energy which is down converted by the nonlinear crystal j in producing the signal and idler photons, α_s (α_i) is the phase shift associated with propagation from NL1 to NL2 along the signal (idler) trajectory, and τ is the time difference between laser light arriving at NL1 and laser light arriving at NL2.

Note that we have dropped the j in equation (5.20) because the s_1 and s_2 and also the i_1 and i_2 trajectories are aligned. If a photon is detected in either of the two output beams then the vacuum term in equation (5.20) becomes redundant and can be dropped. Hence, by choosing

$$f_1 V_1(t) e^{i(\alpha_s + \alpha_i)} = f_2 V_2(t + \tau) \tag{5.21}$$

and adopting our previous notation, the final state is that given in equation (5.18).

If technological advances eventually provide us with detectors with quantum efficiencies greater than 83% then the first direct tests of Bell's inequalities will be possible. Either two photon interference experiments or experiments involving measurements of polarisation with the source proposed above would make such direct tests possible. Unfortunately there is one major practical difficulty that would be encountered in performing the particular polarisation experiment proposed here. This is the problem of aligning the beams s_1 and s_2 and also i_1 and i_2. In the experiment of Zou et al. (1991) where the alignment of the idler beams was necessary in order to see interference between the signal beams, the difficulty of aligning the beams led to only 30% visibility being observed rather than 100% which is theoretically possible. Aligning both beams would pose even greater practical problems and for this reason we will have to wait for experimental techniques to improve sufficiently such that near perfect alignment of the beams is possible before an experimental realisation of this proposal is possible.

5.4 Preparing a singlet state by shuffling boxes

We will now describe a way to prepare two ions in a singlet state by 'shuffling boxes'. This method is probably only possible at the the level of gedanken experiments. However it is an interesting method which can be used to prepare other entangled states as we shall see in later chapters. We will consider ion plus box systems like the one used in chapter 2 in the discussion on empty waves. We take two spin $\frac{1}{2}$ ions, one positively charged (which we will call ion 1) and the other negatively charged (which we will call ion 2) which, when brought together, will combine irreversibly to produce a molecule. Each ion is prepared in the state

$$|\psi_k\rangle = \beta_{k+}|+\rangle_k + \beta_{k-}|-\rangle_k \tag{5.22}$$

where $|\pm\rangle_k$ is the state of particle k when it has spin $\pm\frac{1}{2}$ along the z-axis. Once prepared in this state each ion is placed into a box (box k for ion k) which is longer than it is wide and which is orientated such that its length is along the z-axis (see fig 5.3(a)).

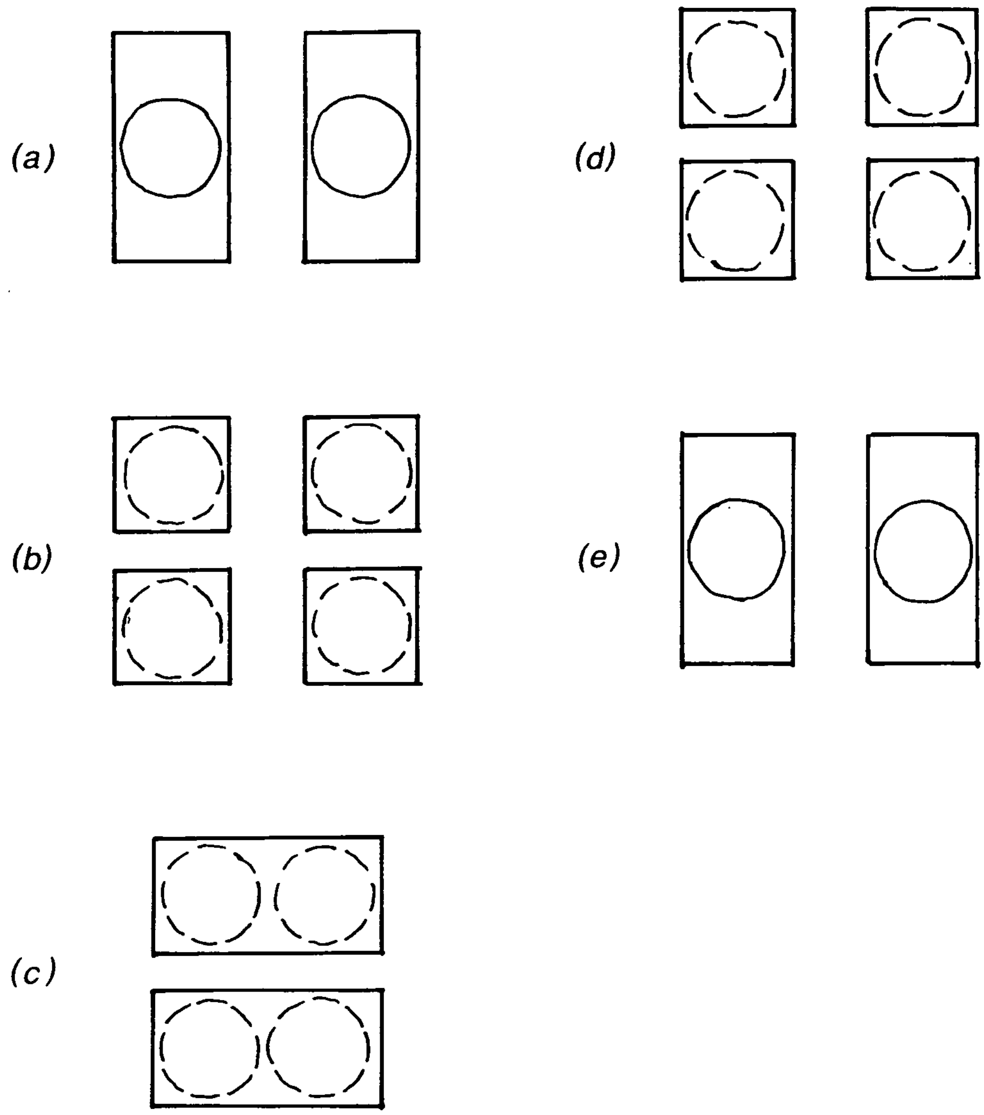

Fig. 5.3. Stages required to produce an entangled state by shuffling boxes.

A nonuniform magnetic field is placed along the length of box k which has the effect of splitting the state of ion k with the $|+\rangle_k$ part going into the upper part of the box and the $|-\rangle_k$ part going into the lower part of the box. Now a partition is placed so as to divide the box into upper and lower boxes (fig. 5.3(b)). The upper box will be called box A_k and the lower box will be called box B_k. The state of each ion plus boxes system is now

$$|\psi_k\rangle = \beta_{k+}|A,+\rangle_k + \beta_{k-}|B,-\rangle_k \tag{5.23}$$

where $|A,\pm\rangle_k$ is the state of the system when ion k is in box A_k and has spin $\pm\frac{1}{2}$ and $|B,\pm\rangle_k$ is the state of the system when ion k is in box B_k and has spin $\pm\frac{1}{2}$. The state of the whole system is simply the product of the states of each system, i.e.

$$\begin{aligned}|\Psi\rangle = \beta_{1+}\beta_{2+}|A,+\rangle_1|A,+\rangle_2 + \beta_{1+}\beta_{2-}|A,+\rangle_1|B,-\rangle_2 \\ +\beta_{1-}\beta_{2+}|B,-\rangle_1|A,+\rangle_2 + \beta_{1-}\beta_{2-}|B,-\rangle_1|B,-\rangle_2\end{aligned} \tag{5.24}$$

Next, boxes A_1 and A_2 are brought together and the wall dividing them is removed to form box A. Likewise, boxes B_1 and B_2 are brought together and the wall dividing them is removed to form box B (fig. 5.3(c)). Now, if both ions are in the same box then they will combine to produce a molecule. That is, we have

$$|A,\pm\rangle_1|A,\pm\rangle_2 \longrightarrow |A, M(\pm,\pm)\rangle \tag{5.25}$$

and

$$|B,\pm\rangle_1|B,\pm\rangle_2 \longrightarrow |B, M(\pm,\pm)\rangle \tag{5.26}$$

where $|A, M(\pm,\pm)\rangle$ ($|B, M(\pm,\pm)\rangle$) is the state of box A (B) when it has a molecule in it that is produced by combining ion 1 with spin $\pm\frac{1}{2}$ and ion 2 with spin $\pm\frac{1}{2}$. A molecule will be formed in half of the runs of the experiment. However, we are interested in those runs of the experiment in which a molecule is not formed. Thus, we imagine that the walls of the box are transparent to neutral molecules and consequently, the molecule will fall out of the bottom of the box under gravity. Next, we apply an electric field across each box such that if there is a positive (negative) ion in the box then it will go to the left hand (right hand) side of the box and then the dividing walls are replaced such that we have four boxes again (fig. 5.3(d)). The electric field has had the effect of returning ion 1 which is positively charged to box A_1 or box B_1 and ion 2 which is negatively charged to box A_2 or box B_2. Thus the second and the third term in (5.24) remain unchanged. However, the first and last term in (5.24) are altered, effectively taken out, by the interactions in (5.25) and (5.26) respectively leaving the state of the system as

$$|\Psi\rangle = \beta_{1+}\beta_{2-}|A,+\rangle_1|B,-\rangle_2 + \beta_{1-}\beta_{2+}|B,-\rangle_1|A,+\rangle_2 + \gamma|\text{molecule}\rangle \tag{5.27}$$

where $|\text{molecule}\rangle$ is the term corresponding to the state of the molecule. The molecule term becomes redundant if future measurements reveal that there are ions in the boxes and consequently we can ignore this term. If we bring box A_k and B_k together (for $k = 1, 2$) and remove the dividing wall (fig 5.3(e)) then the state of the two ions becomes

$$|\Psi\rangle = \omega\Big(\beta_{1+}\beta_{2-}|+\rangle_1|-\rangle_2 + \beta_{1-}\beta_{2+}|-\rangle_1|+\rangle_2\Big) \tag{5.28}$$

where ω is a normalising constant. By putting $\beta_{1+}\beta_{2-} = -\beta_{1-}\beta_{2+}$ (5.28) becomes the singlet state. This could be done by initially preparing ion 1 with spin $+\frac{1}{2}$ along the x-axis and ion 2 with spin $-\frac{1}{2}$ along the same axis.

The technology to trap single charged ions does exist Chu (1992) although it is not clear whether the trapping fields would couple to the spin of the ions and destroy any superposition like that in (5.28). Furthermore, no experiments involving the sort of complex manoeuvres described above have been performed. For these reasons, it seems unlikely that 'shuffling boxes' could become a viable experimental method for the preparation of entangled states. Nevertheless, it remains an interesting method from the gedanken point of view.

Chapter 6

The Nonlocality of a Single Photon

6.1 Introduction

All the experiments to test local realism discussed so far have involved a source that emits pairs of particles, the two particles separating to impinge on two distant detector apparatuses. When a single particle passes through a beam splitter, its state becomes

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\Big(i|1\rangle_1|0\rangle_2 + |0\rangle_1|1\rangle_2\Big) \tag{6.1}$$

This is an entangled state and therefore it seems likely that a demonstration of quantum nonlocality is possible by making some sort of measurements on each of the two beams emerging from this beam splitter.

The first demonstration of this type was due to Tan, Walls, and Collett (1991). We will find, however, that their proposal requires the use of the supplementary assumption even in the case of an ideal experiment. This is a rather serious weakness of the proposal. We will then consider another proposal to demonstrate the nonlocality of a single photon (Hardy (1991b) and (1991c)). This time we find that although this proposal does not suffer from the same weakness as that of TWC it does have another perhaps more serious weakness if it is really to be considered as a demonstration of the nonlocality of a single photon. Finally, we will show how these two proposals can be combined to generate a proposal which is free of both problems.

6.2 The Proposal of Tan, Walls, and Collett

The apparatus considered by TWC is shown in fig. 6.1. A single photon impinges onto a 50–50 beam splitter, $BS3$, with the vacuum incident on the other input and homodyne detection is performed on each of the two outputs from this beam splitter. Each homodyne detector (labelled by $k = 1, 2$) consists of a 50–50 beam splitter BSk with a detector in each of its two outputs c_k and d_k and a coherent local oscillator

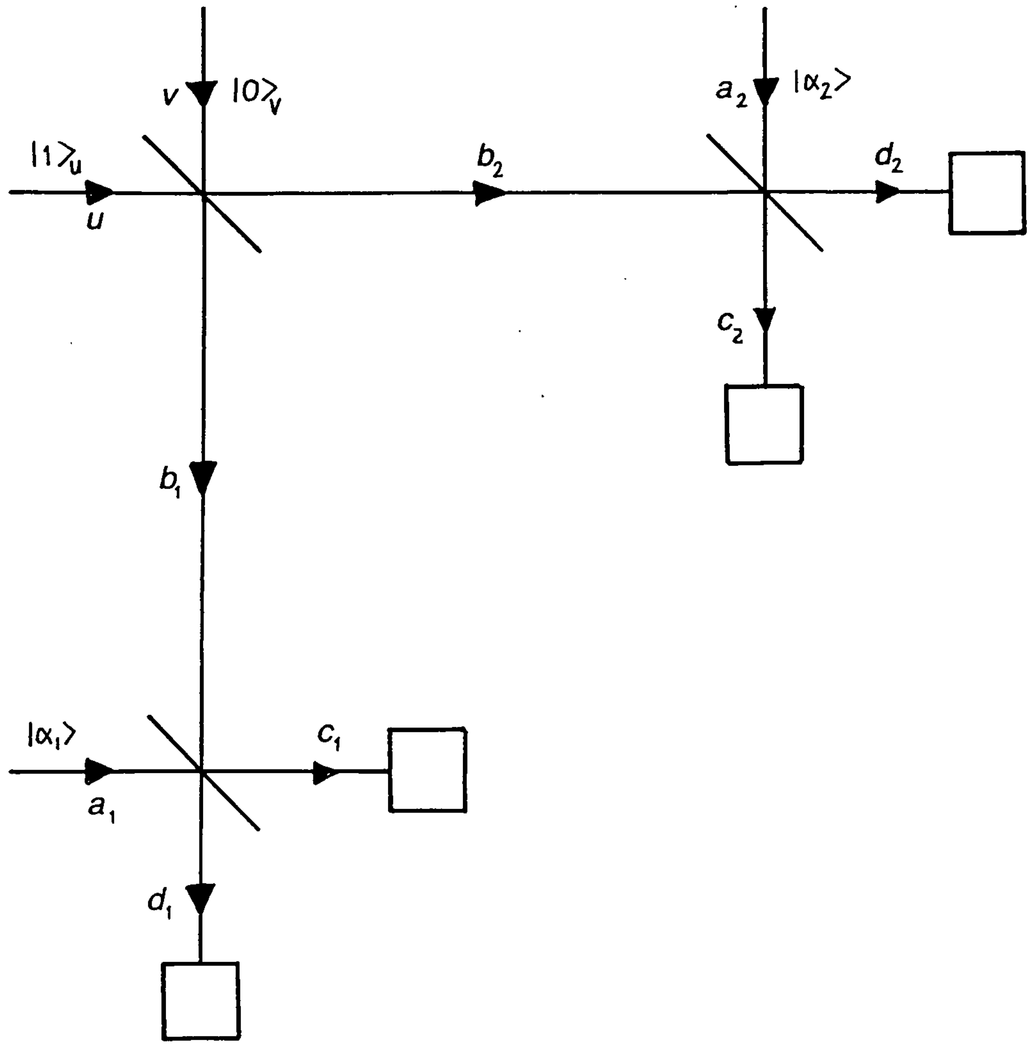

Fig. 6.1. A single photon impinges on a beam splitter and homodyne detection is performed on each output.

with amplitude $\alpha_k = \alpha \exp(i\theta_k)$ incident on one of its inputs. In fact, a very similar experiment has been proposed by Oliver and Stroud (1989) but they did not present it as an experiment to demonstrate the nonlocality of a single photon as did TWC.

At $BS3$ the u and v inputs are transformed into the b_1 and b_2 outputs. The corresponding annihilation (or mode) operators are related by the transformation

$$\begin{pmatrix} \hat{b}_1 \\ \hat{b}_2 \end{pmatrix} = \begin{pmatrix} \sqrt{T} & i\sqrt{R} \\ i\sqrt{R} & \sqrt{T} \end{pmatrix} \begin{pmatrix} \hat{v} \\ \hat{u} \end{pmatrix} \tag{6.2}$$

Likewise, at beam splitter BSk (for $k = 1, 2$) we have the transformation

$$\begin{pmatrix} \hat{c}_k \\ \hat{d}_k \end{pmatrix} = \begin{pmatrix} \sqrt{T} & i\sqrt{R} \\ i\sqrt{R} & \sqrt{T} \end{pmatrix} \begin{pmatrix} \hat{a}_k \\ \hat{b}_k \end{pmatrix} \tag{6.3}$$

(Campos, Saleh and Teich (1989)). The initial state of the system is

$$|1\rangle_u|0\rangle_v|\alpha_1\rangle|\alpha_2\rangle_2 \tag{6.4}$$

which can be written

$$\widehat{u}^\dagger|0\rangle_u|0\rangle_v|\alpha_1\rangle_{a_1}|\alpha_2\rangle_{a_2} \tag{6.5}$$

where $|\alpha_k\rangle_{a_k}$ is the state of the coherent local oscillator in mode a_k. These coherent states are eigenstates of the annihilation operator with eigenvalue equation

$$\widehat{a}_k|\alpha_k\rangle = \alpha_k|\alpha_k\rangle \tag{6.6}$$

From (6.2) we have

$$\widehat{u}^\dagger \longrightarrow \frac{1}{\sqrt{2}}(i\widehat{b}_1^\dagger + \widehat{b}_2^\dagger) \tag{6.7}$$

Thus, on evolving through beam splitter $BS3$ the state becomes

$$\frac{1}{\sqrt{2}}(i\widehat{b}_1^\dagger + \widehat{b}_2^\dagger)|0\rangle_{b_1}|0\rangle_{b_2}|\alpha_1\rangle|\alpha_2\rangle_2 \tag{6.8}$$

which we will write as

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\Big(i|\varphi_1\rangle + |\varphi_2\rangle\Big) \tag{6.9}$$

with

$$|\varphi_1\rangle = |1\rangle_{b_1}|0\rangle_{b_2}|\alpha_1\rangle_{a_2}|\alpha_2\rangle_{a_2} \tag{6.10}$$

and

$$|\varphi_2\rangle = |0\rangle_{b_1}|1\rangle_{b_2}|\alpha_1\rangle_{a_2}|\alpha_2\rangle_{a_2} \tag{6.11}$$

When the state evolves through the beam splitters BSk it becomes rather more cumbersome to write down and so instead we will perform our calculations with the state in (6.9) and use the operator transformations in (6.3). The intensity operator for the d_1 output mode is $\widehat{I}_{d_1} = \widehat{d}_1^\dagger\widehat{d}_1$. Thus the expectation value of the intensity in the output mode d_1, that is given by finding the expectation value of $\frac{1}{2}(-i\widehat{a}_1^\dagger + \widehat{b}_1^\dagger)(i\widehat{a}_1 + \widehat{b}_1)$ evaluated with the state (6.9). We are interested in correlations in counts between detectors at each end of the apparatus, that is the expectation values of quantities like $I_{d_1}I_{d_2}$. When evaluated in terms of $\widehat{a}_k$ and $\widehat{b}_k$ using (6.3), $\widehat{I}_{d_1}\widehat{I}_{d_2}$ will consist of sixteen terms. Fortunately calculations are made simpler by the fact that only a few terms actually contribute to the expectation value.

Using these methods it can be shown that

$$\langle\varphi_k|I_{d_1}I_{d_2}|\varphi_k\rangle = \langle\varphi_k|I_{d_1}I_{c_2}|\varphi_k\rangle = \langle\varphi_k|I_{c_1}I_{d_2}|\varphi_k\rangle = \langle\varphi_k|I_{c_1}I_{c_2}|\varphi_k\rangle = \frac{1}{4}(\alpha^4+\alpha^2) \quad (6.12)$$

In this case the $\widehat{b}_1^\dagger\widehat{a}_2^\dagger\widehat{b}_1\widehat{a}_2$ term gives rise to the α^4 term and the $\widehat{a}_1^\dagger\widehat{a}_2^\dagger\widehat{a}_1\widehat{a}_2$ term in the case of $k=1$ or the $\widehat{a}_1^\dagger\widehat{b}_2^\dagger\widehat{a}_1\widehat{b}_2$ in the case of $k=2$ gives rise to the α^2 term. This is quite easily seen from equations (6.6), (6.10) and (6.11). The cross terms can also be evaluated giving

$$\langle\varphi_2|I_{d_1}I_{d_2}|\varphi_1\rangle = \langle\varphi_2|I_{c_1}I_{c_2}|\varphi_1\rangle = \frac{1}{4}(\alpha^4+\alpha_1^*\alpha_2) \quad (6.13)$$

$$\langle\varphi_2|I_{d_1}I_{c_2}|\varphi_1\rangle = \langle\varphi_2|I_{c_1}I_{d_2}|\varphi_1\rangle = \frac{1}{4}(\alpha^4-\alpha_1^*\alpha_2) \quad (6.14)$$

As before the $\widehat{b}_1^\dagger\widehat{a}_2^\dagger\widehat{b}_1\widehat{a}_2$ term gives rise to the α^4 terms. The $\alpha_1^*\alpha_2$ terms are due to the $\widehat{a}_1^\dagger\widehat{b}_2^\dagger\widehat{b}_1\widehat{a}_2$ term.

Using (6.9), (6.12), (6.13) and (6.14) we obtain the expectation values

$$\langle I_{d_1}I_{d_2}\rangle = \langle I_{c_1}I_{c_2}\rangle = \frac{1}{4}\Big(\alpha^4+\alpha^2\big(1+\sin(\theta_1-\theta_2)\big)\Big) \quad (6.15)$$

$$\langle I_{d_1}I_{c_2}\rangle = \langle I_{d_1}I_{c_2}\rangle = \frac{1}{4}\Big(\alpha^4+\alpha^2\big(1-\sin(\theta_1-\theta_2)\big)\Big) \quad (6.16)$$

TWC considered the correlation function

$$E(\theta_1,\theta_2) = \frac{\langle(I_{d_1}-I_{c_1})(I_{d_2}-I_{c_2})\rangle}{\langle(I_{d_1}+I_{c_1})(I_{d_2}+I_{c_2})\rangle} \quad (6.17)$$

Using (6.15) and (6.16) we obtain

$$E(\theta_1,\theta_2) = \frac{1}{\alpha^2+1}\sin(\theta_1-\theta_2) \quad (6.18)$$

TWC then note that if we have a correlation function of the form in (6.18) with a coefficient greater than $\frac{1}{\sqrt{2}}$ in front of a $\sin(\theta_1-\theta_2)$ function then Bell's inequalities are violated. Thus by choosing α to be sufficiently small the correlation function (6.18) will violate Bell's inequalities demonstrating the nonlocality of a single photon.

However, the correlation function (6.17) is not the same as the correlation function considered in chapter 4 that enters into the CHSH Bell inequalities. It is more similar to the correlation function in section 5.2 used in the discussion on the supplementary assumption except that there, the correlation function was defined in terms of probabilities rather than intensities that can take values greater than 1. In section 5.2 we saw that the supplementary assumption was not necessary in ideal experiments. However, as we will now show, in the proposal of TWC, the assumption is needed *even in an ideal experiment* in which we have detectors that are capable of measuring the intensity (i.e. the number of photons) of the incident beam with 100% accuracy. The supplementary assumption required is

$$I_{\widehat{d_1}}(\lambda, \theta_1) + I_{\widehat{c_1}}(\lambda, \theta_1) + I_{\widehat{d_2}}(\lambda, \theta_2) + I_{\widehat{c_2}}(\lambda, \theta_2) = I(\lambda) \tag{6.19}$$

The interpretation of this assumption is as follows: For every λ, the total intensity emerging from the beam splitters at ends 1 and 2 of the apparatus is independent of the settings of the local variables θ_k. If θ_l is kept constant then I must be insensitive to variation in θ_k $(k \neq l)$. Consequently,

$$I_{\widehat{d_k}}(\lambda, \theta_k) + I_{\widehat{c_k}}(\lambda, \theta_k) = I_k(\lambda) \tag{6.20}$$

follows from (6.19). Using (6.20) it is possible to derive Bell inequalities with the same form as the CHSH inequalities but which employ the correlation coefficient (6.17) by putting

$$x(\lambda) = I_{\widehat{d_1}}(\lambda, \theta_1) - I_{\widehat{c_1}}(\lambda, \theta_1) \tag{6.21}$$

$$x'(\lambda) = I_{\widehat{d_1}}(\lambda, \theta_1') - I_{\widehat{c_1}}(\lambda, \theta_1') \tag{6.22}$$

$$y(\lambda) = I_{\widehat{d_2}}(\lambda, \theta_2) - I_{\widehat{c_2}}(\lambda, \theta_2) \tag{6.23}$$

$$y(\lambda) = I_{\widehat{d_2}}(\lambda, \theta_2') - I_{\widehat{c_2}}(\lambda, \theta_2') \tag{6.24}$$

$$X(\lambda) = I_1(\lambda) \tag{6.25}$$

$$Y(\lambda) = I_2(\lambda) \tag{6.26}$$

into inequalities (5.5). (Another way of deriving these inequalities is given by Żukowski (1989).) The choices (6.21) to (6.25) satisfy the inequalities (5.1) and (5.2) used in deriving the inequalities (5.5) if (6.20) is true. The CHSH type Bell inequalities derived in this way are violated by the correlation function (6.18) for appropriate choices of local variable setting (see fig. 4.2) if $1/(1+\alpha^2) > 1/\sqrt{2}$.

In standard experiments to test Bell's inequalities $I(\lambda) = 2$ for all λ (because there are two particles) such that the RHS of (6.19) is constant and independent of λ. Now, in an ideal experiment this could be verified experimentally by measuring I and noticing that $I = 2$ for every run of the experiment. Therefore, in an ideal experiment of the standard type, (6.19) is not an assumption. On the other hand, if we use a homodyne detection scheme as in the proposal of TWC (see also Grangier et al. (1988)) where, at each end of the apparatus, the beam from the source is mixed at a beam splitter with a beam from a coherent local oscillator, then the total intensity I is not a constant because the beams from the coherent local oscillators are not in a number state. Therefore, the RHS of (6.19) necessarily depends on λ and, consequently, (6.19) cannot be verified by experiment even in an ideal experiment. Thus, even in an ideal experiment, the supplementary assumption must be used.

One final point, the assumption (6.19) must strictly be interpreted as conservation of intensity under variation of θ_k. It is not a conservation of probability equation like equation (5.9). However, as modes $\widehat{c}_k$ and $\widehat{d}_k$ are not in number states, such notions of conservation of intensity cannot necessarily be regarded as quantum mechanical concepts. Therefore, we should not be surprised if assumption (6.19) is not true in the proposal of TWC. This experiments are, then, less valuable in demonstrating nonlocality in quantum theory than standard experiments. It is nevertheless still of interest because it addresses the question of whether a single photon can cause a nonlocal correlation.

6.3 Nonlocality using a Mach-Zehnder interferometer

In this section we consider a different way of demonstrating the nonlocality of a single photon. The proposal here is motivated by the Gedanken experiment of Elitzur and Vaidman (1991) discussed in chapter 3. Consider a Mach-Zehnder interferometer like that shown in fig. 6.2 with a single photon incident in the u input. The vacuum is incident on the other input. The interferometer has two internal paths b_1 and b_2 and two outputs c and d. The beam splitters $BS1$ and $BS2$ are 50–50. The path lengths in the interferometer are set such that, when there are no objects in the internal paths, no photons will emerge at output d due to destructive interference. Next, we take two spin $\frac{1}{2}$ atoms. Each is prepared in the state

$$|\psi_k\rangle = \beta_{k+}|+\rangle_k + \beta_{k-}|-\rangle_k \tag{6.27}$$

where $|\pm\rangle_k$ is the state of atom k when it has spin $\pm\frac{1}{2}$ along the z-axis. Each atom, prepared in this state is then placed in a box and in the same way as described in sections 3.5 and 5.4, a nonuniform magnetic field is placed along the length of the box and then and then a wall is placed so as to divide the box into two , box A_k with the spin up part of the state and box B_k with the spin down part. The state of each atom plus box system is

$$|\psi_k\rangle = \beta_{k+}|A,+\rangle_k + \beta_{k-}|B,-\rangle_k \tag{6.28}$$

where we are using the same notation as in section 5.4. We will assume that the walls of the boxes are transparent to the photons but that if the atom is in the path of the photon then it will absorb the photon with probability equal to 1. Thus we have

$$|1\rangle_{b_k}|A,+\rangle_k \longrightarrow |0\rangle_{b_k}|A,+\rangle_k^{ex} \tag{6.29}$$

Now the atom 1 plus box system is placed with box A_1 in path b_1 and likewise the atom 2 plus box system is placed so that box A_2 is in path b_2. We will be interested in those runs of the experiment for which the photon is detected in the previously dark output d. When this happens one of the two paths through the interferometer must be blocked so that there is no longer destructive interference at output d and one path must be open otherwise the photon would have been blocked from going through the interferometer. There are two ways in which this could happen. Either atom 1 is in box A_1 and atom 2 is in box B_2 or atom 1 is in box B_1 and atom 2 is in box A_2. Thus the state of the atoms, when a photon is detected in the d output must be

$$\mu|A,+\rangle_1|B,-\rangle_2 + \nu|B,-\rangle_1|A,+\rangle_2 \tag{6.30}$$

We will now see in detail how this result arises.

The operation of beam splitter $BS1$ is, using (6.2),

$$|1\rangle_u|0\rangle_v \longrightarrow \frac{1}{\sqrt{2}}\Big(i|1\rangle_{b_1}|0\rangle_{b_2} + |0\rangle_{b_1}|1\rangle_{b_2}\Big) \tag{6.31}$$

and the operation of beam splitter $BS2$ is, using (6.3),

$$|1\rangle_{b_1}|0\rangle_{b_2} \longrightarrow \frac{1}{\sqrt{2}}\Big(|1\rangle_c|0\rangle_d + i|0\rangle_c|1\rangle_d\Big) \tag{6.32}$$

$$|0\rangle_{b_1}|1\rangle_{b_2} \longrightarrow \frac{1}{\sqrt{2}}\Big(i|1\rangle_c|0\rangle_d + |0\rangle_c|1\rangle_d\Big) \tag{6.33}$$

The initial state of the system is

$$|1\rangle_u|0\rangle_v|\psi_1\rangle|\psi_2\rangle \tag{6.34}$$

After the photon has passed through $BS1$, but before it reaches the boxes, the state of

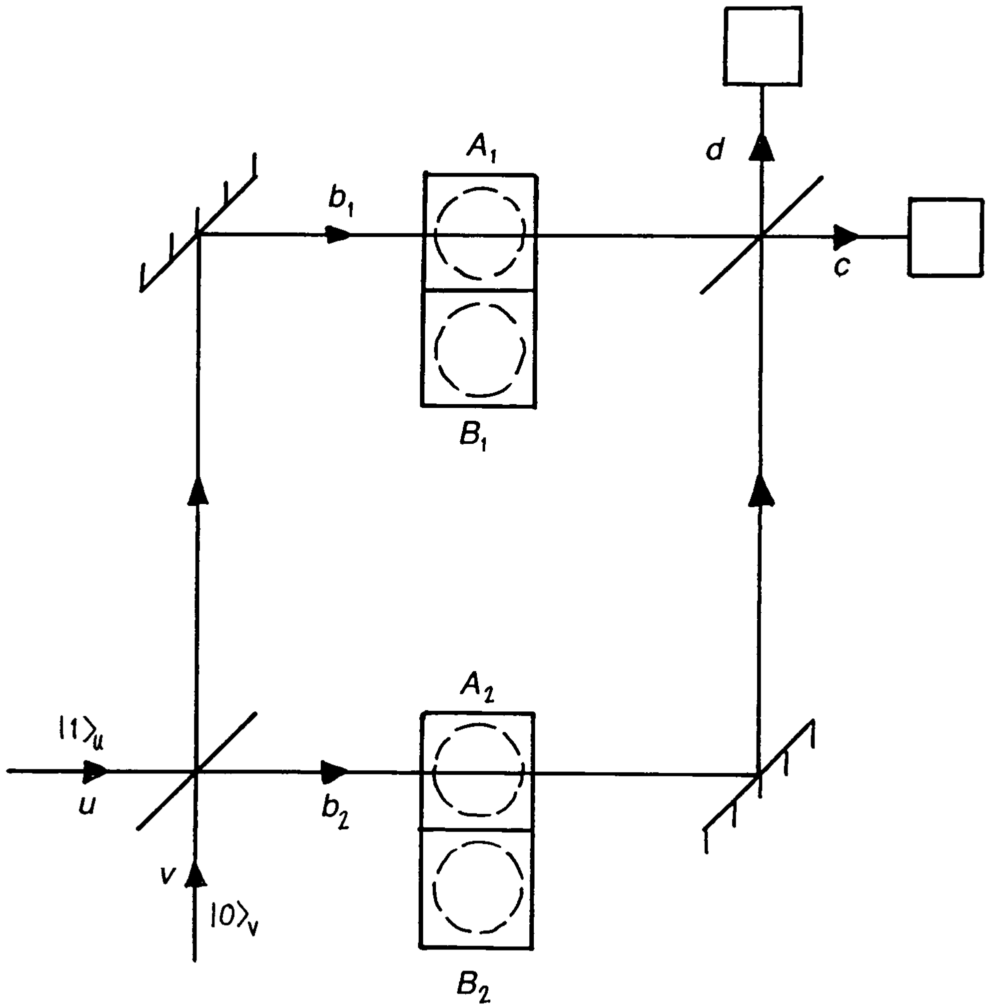

Fig. 6.2. A Mach-Zehnder interferometer with an 'atom plus boxes system' placed in each internal path.

the system is

$$\frac{1}{\sqrt{2}}\Big(i|1\rangle_{b_1}|0\rangle_{b_2} + |0\rangle_{b_1}|1\rangle_{b_2}\Big)|\psi_1\rangle|\psi_2\rangle \tag{6.35}$$

After the photon interacts with the atoms the state of the system becomes, using (6.28) and (6.29),

$$\begin{aligned}&\frac{1}{\sqrt{2}}\Big(i\beta_{1+}|0\rangle_{b_1}|0\rangle_{b_2}|A,+\rangle_1^{ex} + i\beta_{1-}|0\rangle_{b_1}|1\rangle_{b_2}|B,-\rangle_1\Big)|\psi_2\rangle \\ &\quad + \frac{1}{\sqrt{2}}\Big(\beta_{2+}|0\rangle_{b_1}|0\rangle_{b_2}|A,+\rangle_2^{ex} + \beta_{2-}|0\rangle_{b_1}|1\rangle_{b_2}|B,-\rangle_2\Big)|\psi_2\rangle\end{aligned} \tag{6.36}$$

After evolving through $BS2$ the state becomes

$$\frac{1}{\sqrt{2}}\Big(i\beta_{1+}|0\rangle_c|0\rangle_d|A,+\rangle_1^{ex}|\psi_2\rangle + \beta_{2+}|0\rangle_c|0\rangle_d|A,+\rangle_2^{ex}|\psi_1\rangle\Big)$$
$$+\frac{i}{2}|1\rangle_c|0\rangle_d\Big(\beta_{1-}|B,-\rangle_1|\psi_2\rangle + \beta_{2-}|B,-\rangle_2|\psi_1\rangle\Big) \tag{6.37}$$
$$+\frac{1}{2}|0\rangle_c|1\rangle_d\Big(\beta_{1+}\beta_{2-}|A,+\rangle_1|B,-\rangle_2 - \beta_{1-}\beta_{2+}|B,-\rangle_1|A,+\rangle_2\Big)$$

If the photon is detected in output d then the state of the system is projected onto the last term in (6.37) and consequently if the dividing walls in the boxes are removed then the (unnormalized) state of the atoms becomes

$$|\Phi\rangle = \beta_{1+}\beta_{2-}|+\rangle_1|-\rangle_2 - \beta_{1-}\beta_{2+}|-\rangle_1|+\rangle_2 \tag{6.38}$$

If the two atoms are initially prepared in the same state so that $\beta_{1\pm} = \beta_{2\pm}$ then the state (6.38) is the singlet state. For this state we know that measurements of spin along appropriate directions will violate the CHSH inequalities.

The experiment proceeds in the following way. First the photon is be sent through the apparatus. If the photon is detected in the dark output then measurements of spin along some directions $\mathbf{a}$ on atom 1 and $\mathbf{b}$ on atom 2 (see section 4.3) are be performed. The expectation value of the product of these spin measurements is then be taken. Thus, the actual quantity being measured is

$$E(\mathbf{a},\mathbf{b}) = \frac{\langle\sigma_{1\mathbf{a}}\sigma_{2\mathbf{b}}I_d\rangle}{\langle I_d\rangle} \tag{6.39}$$

where $I_d = \hat{d}^\dagger\hat{d}$ is the number operator for the d output. This correlation function is not of the form considered in chapter 4 and therefore it is not clear that it can be used in Bell inequalities. To see that it can let $\sigma_{1\mathbf{a}}(\lambda)$, $\sigma_{2\mathbf{b}}(\lambda)$, and $I_d(\lambda)$ be the actual values of the quantities. We note that, as the choice of measurements of spin direction can be made after the photon has passed the atoms, the quantity I_d does not depend on $\mathbf{a}$ or $\mathbf{b}$. Therefore we can put

$$x(\lambda) = \sigma_{1\mathbf{a}}(\lambda) \tag{6.40}$$

$$x'(\lambda) = \sigma_{1\mathbf{a}'}(\lambda) \tag{6.41}$$

$$y(\lambda) = \sigma_{2\mathbf{b}}(\lambda)I_d(\lambda) \tag{6.42}$$

$$y'(\lambda) = \sigma_{2\mathbf{b}'}(\lambda)I_d(\lambda) \tag{6.43}$$

$$X(\lambda) = 1 \tag{6.44}$$

$$Y(\lambda) = I_d(\lambda) \tag{6.45}$$

for the quantities in section 5.2. These choices satisfy inequalities (5.1) and (5.2) without the need for any supplementary assumptions. By substituting (6.40) to (6.45) into inequalities (5.5) we obtain CHSH-type Bell inequalities that employ the correlation function (6.39).

This proposal does not suffer from the problem that the proposal of TWC does, that is it does not require a supplementary assumption. However, it could be argued that it does not demonstrate the nonlocality of a single photon because the two paths of the photon are brought back together. Thus, although only one photon emerges from the source, this photon will be influenced by the hidden variables associated with the atoms and this will influence which way it goes at $BS2$. Therefore, the particular subensemble that we consider (i.e. the one for which there is a detection in output d) is determined by the hidden variables associated with three particles. It would be far more satisfactory if the two paths of the photon were not brought back together.

6.4 The nonlocality of a single photon without problems

Neither of the two proposals we have considered so far demonstrate the nonlocality of a single photon in an entirely satisfactory way. In the proposal of TWC there is the need for a supplementary assumption even in the case of an ideal experiment and in the proposal using a Mach-Zehnder interferometer the two paths of the photon are brought back together. We will now see that by combining these two experiments we can avoid both problems. The apparatus we will consider is shown in fig 6.3. This is the same as the apparatus of TWC except that an atom plus box system has been placed with box A_k in each of the paths b_k like in the second proposal above. The phases θ_k are set so that $\sin(\theta_1 - \theta_2) = -1$ and hence, with the boxes removed from the paths, $\langle I_{d_1} I_{d_2} \rangle = \alpha^4$ from (6.15). Therefore, if α is small, the probability of a coincidence count between the detectors in d_1 and d_2 is very small for these values of θ_k. Thus we can consider a $d_1 d_2$ coincidence count to be analogous to a count in the dark output d of the Mach-Zehnder interferometer in the second proposal above. If the boxes are now put into the paths (as shown in fig 6.3) and there is a $d_1 d_2$ coincidence detection then by, comparison with the Mach-Zehnder proposal, we would expect the atoms to go into an entangled state like that in (6.38). We will find that this is the case although it is a little more involved than in the Mach-Zehnder proposal.

The initial state of the system is

$$|1\rangle_u |0\rangle_v |\alpha_1\rangle |\alpha_2\rangle |\psi_1\rangle || \psi_2\rangle \tag{6.46}$$

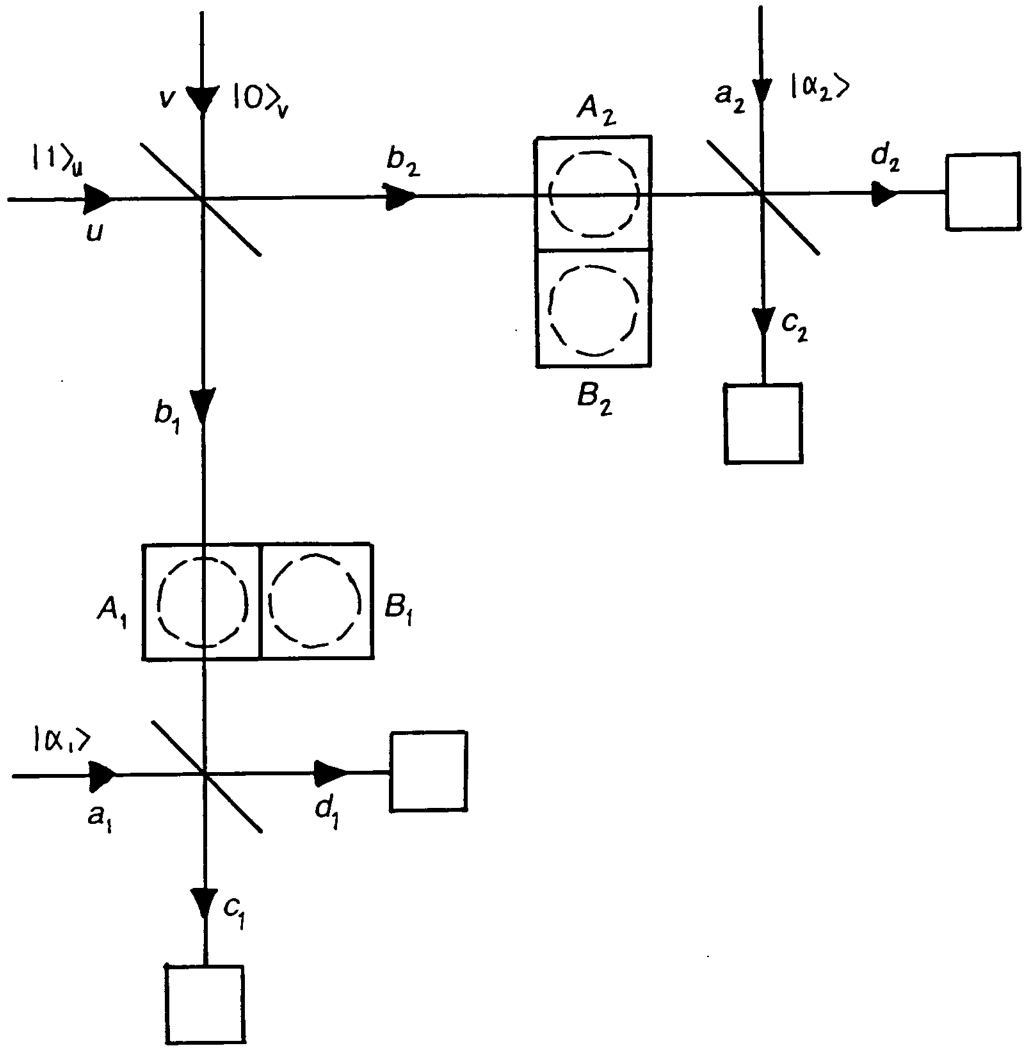

Fig. 6.3. The apparatus of Tan, Walls and Collett with an 'atom plus boxes system' placed in each internal path.

After the photon passes through $BS2$ the state becomes

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\Big(i|\varphi_1\rangle + |\varphi_2\rangle\Big)|\psi_1\rangle|\psi_2\rangle \tag{6.47}$$

(cf. (6.9)). From (6.10), (6.11) and (6.29) the interaction between the photon and each atom is given by

$$|\varphi_k\rangle|A,+\rangle_k \longrightarrow |\varphi, A, +\rangle_{kex} \tag{6.48}$$

where

$$|\varphi, A, +\rangle_k^{ex} = |0\rangle_{b_1}|0\rangle_{b_2}|A,+\rangle_k^{ex}|\alpha_1\rangle_{a_2}|\alpha_2\rangle_{a_2} \tag{6.49}$$

Thus after the interaction between the photon and the atoms, the state becomes, using

(6.28) and (6.48),

$$\frac{1}{\sqrt{2}}\Big(i\beta_{1+}|\varphi, A, +\rangle_1^{ex}|\psi_2\rangle + i\beta_{2+}|\varphi, A, +\rangle_2^{ex}|\psi_1\rangle\beta_{1-}|\varphi_1\rangle|B, -\rangle_1|\psi_2\rangle + \beta_{2-}|\varphi_2\rangle|B, -\rangle_2|\psi_1\rangle\Big) \tag{6.50}$$

It is now possible to make a measurement on each of the atoms to see whether they are excited. A negative result for this measurement projects the system into the (unnormalized) state,

$$\Big(\beta_{1-}|\varphi_1\rangle|-\rangle_1|\psi_2\rangle + \beta_{2-}|\varphi_2\rangle|-\rangle_2|\psi_1\rangle\Big) \tag{6.51}$$

where the dividing walls in the boxes have now been removed. Note that the measurement on the atoms to see if they are excited must not disturb the other possible states of the atoms. Such a measurement could be performed by surrounding both box A_1 and box A_2 by detectors to detect the photon when the atom which had absorbed it de-excites. If, after a time much greater than the de-excitation time, there has been no such detection then the system is projected into the state in (6.51).

The experiment proceeds in the following way. in the following way. If neither atom is excited and if there is a $d_1 d_2$ coincidence count then measurements of spin along directions **a** and **b** are made on atoms 1 and 2 respectively. Thus the quantity actually measured is

$$E(\mathrm{a}, \mathrm{b}) = \frac{\langle \sigma_{1\mathrm{a}}\sigma_{2\mathrm{b}} I_{d_1} I_{d_2} P_1 P_2\rangle}{\langle I_{d_1} I_{d_2} P_1 P_2\rangle} \tag{6.52}$$

where P_k is equal to 1 if atom k is not excited and equal to 0 if atom k is excited. A Bell inequality can be formed from the correlation function in (6.52) by putting

$$x(\lambda) = \sigma_{1\mathrm{a}}(\lambda) I_{d_1}(\lambda) P_1(\lambda) \tag{6.53}$$

$$x'(\lambda) = \sigma_{1\mathrm{a}'}(\lambda) I_{d_1}(\lambda) P_1(\lambda) \tag{6.54}$$

$$y(\lambda) = \sigma_{2\mathrm{b}}(\lambda) I_{d_2}(\lambda) P_2(\lambda) \tag{6.55}$$

$$y'(\lambda) = \sigma_{2\mathrm{b}'}(\lambda) I_{d_2}(\lambda) P_2(\lambda) \tag{6.56}$$

$$X(\lambda) = I_{d_1}(\lambda) P_1(\lambda) \tag{6.57}$$

$$Y(\lambda) = I_{d_2}(\lambda) P_2(\lambda) \tag{6.58}$$

in inequality (5.5). The Bell inequality formed in this way has the same form as the CHSH inequalities. The quantities $I_{d_k}(\lambda)$ and $P_k(\lambda)$ do not depend on the settings of **a** and **b** because these settings are made after the measurements I_{d_k} and P_k and consequently the choices (6.53) to (6.58) satisfy inequalities (5.1) and (5.2) without the need for a supplementary assumption.

It is now a simple matter to calculate $E(\mathbf{a},\mathbf{b})$. By using the projected state (6.51) we take care of the P_1P_2 factor in (6.52). The numerator of the RHS of (6.52) is

$$\begin{aligned}\langle\sigma_{1\mathbf{a}}\sigma_{2\mathbf{b}}I_{d_1}I_{d_2}P_1P_2\rangle = &+\beta_{1-}^*\beta_{1-1}\langle -|\langle\psi_2|\sigma_{1\mathbf{a}}\sigma_{2\mathbf{b}}|-\rangle_1|\psi_2\rangle\langle\varphi_1|I_{d_1}I_{d_2}|\varphi_1\rangle \\ &- i\beta_{1-}^*\beta_{2-1}\langle -|\langle\psi_2|\sigma_{1\mathbf{a}}\sigma_{2\mathbf{b}}|-\rangle_2|\psi_1\rangle\langle\varphi_1|I_{d_1}I_{d_2}|\varphi_2\rangle \\ &+ i\beta_{2-}^*\beta_{1-2}\langle -|\langle\psi_1|\sigma_{1\mathbf{a}}\sigma_{2\mathbf{b}}|-\rangle_1|\psi_2\rangle\langle\varphi_2|I_{d_1}I_{d_2}|\varphi_1\rangle \\ &+ i\beta_{2-}^*\beta_{2-2}\langle -|\langle\psi_1|\sigma_{1\mathbf{a}}\sigma_{2\mathbf{b}}|-\rangle_2|\psi_1\rangle\langle\varphi_2|I_{d_1}I_{d_2}|\varphi_2\rangle\end{aligned} \tag{6.59}$$

Now, the phases θ_k are set such that the number of d_1d_2 coincidences is a minimum. From (6.15) we see that this is achieved when $\theta_1 - \theta_2 = -\frac{\pi}{2}$ which is ensured by putting

$$\alpha_1^*\alpha_2 = i\alpha^2 \tag{6.60}$$

Furthermore the amplitude α of the local oscillators is chosen to be small such that $\alpha^4 \approx 0$ and can be ignored. With this approximation, we find that (6.59) becomes, using (6.12) to (6.14) and (6.60),

$$\frac{\alpha^2}{4}\langle\Phi|\sigma_{1\mathbf{a}}\sigma_{2\mathbf{b}}|\Phi\rangle \tag{6.61}$$

where

$$|\Phi\rangle = {}_-\beta_{1-}|-\rangle_1|\psi_2\rangle + {}_-\beta_{2-}|-\rangle_2|\psi_1\rangle$$

which, by using (6.28), can be simplified to the state in (6.38). The denominator of the RHS of (6.52) can be calculated in the same way to give

$$\langle I_{d_1}I_{d_2}P_1P_2\rangle = \frac{\alpha^2}{4}\langle\Phi||\Phi\rangle \tag{6.62}$$

Thus we have from (6.52), (6.61), and (6.62),

$$E(\mathbf{a},\mathbf{b}) = \frac{\langle\Phi|\sigma_{1\mathbf{a}}\sigma_{2\mathbf{b}}|\Phi\rangle}{\langle\Phi|\Phi\rangle} \tag{6.63}$$

Thus the state of the atoms when $I_{d_1}I_{d_2}P_1P_2 \neq 0$ is the same as the state of the atoms in the Mach-Zehnder proposal when there is a detection in the dark output d, namely that given in (6.38). As we noted before, by putting $\beta_{1\pm} = \beta_{2\pm}$ this state becomes the singlet state. The correlation function (6.63) will violate the CHSH-type Bell inequalities formed by substituting (6.53) to (6.58) into inequalities (5.5).

This method of demonstrating the nonlocality of a single photon does not require the use of a supplementary assumption and also it is not necessary to bring the two paths of the photon together as in the Mach-Zehnder proposal. Therefore, we conclude that a single photon can exhibit nonlocality.

Chapter 7

Bell's Theorem With More Than Two Particles

7.1 Introduction

The demonstration of quantum nonlocality using the singlet state presented in chapter 4 has two shortcomings

(1) Inequalities are required

(2) It does not apply to every run of the experiment.

In the limit in which the number of settings of the local variables considered at each end tends to infinity that was considered in section 4.4 both these problems are overcome. Recall that as $K \to \infty$, where K is the total number of settings of the local variables considered, we found that $P \geq 1$ where P is the probability of getting a contradiction for a given run of the experiment. As probabilities cannot be greater than one we have $P = 1$. If $P = 1$ then shortcoming (2) is overcome. Furthermore, $P = 1$ is not an inequality and in this sense, we can express the contradiction without inequalities. However, in order to obtain $P = 1$ it was necessary to consider a limiting process involving an inequality. Elitzur, Popescu and Rohrlich (1992) have also shown how it is possible to obtain $P = 1$ but their argument also uses inequalities. The shortcomings (1) and (2) are overcome without considering an infinite number of settings and without using inequalities at any stage of the argument in a demonstration of quantum nonlocality due to Greenberger, Horne, and Zeilinger (1989) (see also Greenberger et al. (1990), Clifton, Redhead and Butterfield (1991) and Pagonis, Redhead and Clifton (1991)). Their demonstration is made possible by considering more than two particles. Another inequality-free demonstration of nonlocality due to Heywood and Redhead (1983) employs two spin 1 particles in singlet state. In this demonstration a Bell-Kochen-Specker type argument is used to demonstrate contextuality for spacelike separated measurements (Bell (1966) and Kochen and Specker (1967)). This has been simplified by Stairs (1983) (see also Brown and Svetlichny (1990)).

In this chapter we will first review the GHZ demonstration. In fact we will present a simplification due to Mermin (1990b,c). In a real experiment the conditions necessary to run the argument without inequalities would no longer apply and therefore, it is necessary to derive inequalities that apply to three or more particle situations. We will consider two ways of doing this. The first is a generalisation of the CHSH method considered in section 4.3 and the second is a generalisation of the method considered in section 4.4 which was in fact motivated by the GHZ argument. Other N-particle Bell inequalities have been derived by Mermin (1990a) and by Suppes and Zanotti (1991). Finally, we will propose two new ways of preparing the type of N-particle states considered by GHZ. The first is an extension of the Mach-Zehnder proposal of the previous chapter (section 6.3). The second is an extension of the 'shuffling boxes' proposal in chapter 5 (section 5.4). These proposals are to be considered as gedanken experiments because their actual realisation seems unlikely. Other experiments have been proposed to prepare GHZ correlations: Żukowski (1991) proposed that homodyne detection is performed on each of four outputs from a nonlinear crystal being pumped by a laser to produce pairs of photons. Yurke and Stoler (1992a) have proposed an experiment in which three photons are emitted simultaneously from three independent sources each impinge onto a beam splitter and then the outputs from each beam splitter are mixed at three further beam splitters.

7.2 The Greenberger, Horne, and Zeilinger demonstration

To extend Bell's theorem to N particles we consider the simultaneous measurements $A_n(a_n)$ in the separate regions R_n for $n = 1$ to N. We use the notation $A_1(a_1)$, $A_2(a_2)$, etc. rather than $A(a)$, $B(b)$, etc. to achieve greater notational simplicity. The possible values of $A_n(a_n)$ are ± 1 and 0, where 0 corresponds to a null measurement (when the measurement apparatus has failed to register a result). Hence,

$$|A_n(a_n, \lambda)| \leq 1 \tag{7.1}$$

where $A_n(a_n, \lambda)$ is the value of $A_n(a_n)$ when the hidden variables are λ. If there are no null results then (1) is replaced by

$$A_n(a_n, \lambda) = \pm 1 \tag{7.2}$$

We consider the correlation coefficient,

$$E(a) = \langle \prod_{n=1}^{N} A_n(a_n) \rangle \tag{7.3}$$

where $E(a)$ is shorthand for $E(a_1, a_2, ...a_N)$. The angle brackets, $\langle \quad \rangle$, denote the average over an infinite number of experiments. For local theories it is possible to write,

$$E(a) = \int \left(\prod_{n=1}^{N} A_n(a_n, \lambda) \right) \rho(\lambda) d\lambda \tag{7.4}$$

where $\rho(\lambda)$ is the probability distribution function of λ.

The GHZ demonstration is very simple. We will consider the simplification of their result due to Mermin (1990b) for three particles (see also Stapp (1990)). Consider the three spin-$\frac{1}{2}$ particle state

$$|\Psi\rangle \frac{1}{\sqrt{2}}(|+\rangle_1|+\rangle_2|+\rangle_3 - |-\rangle_1|-\rangle_2|-\rangle_3) \tag{7.5}$$

where $|\pm\rangle_n$ represents the state of particle n with spin $\pm\frac{1}{2}$ along the z-axis. We now consider measurements of spin along the x and y axes. Quantum mechanics predicts

$$E(x, y, y) = \langle \sigma_x^1 \sigma_y^2 \sigma_y^3 \rangle = +1 \tag{7.6}$$

$$E(x, x, x) = \langle \sigma_x^1 \sigma_x^2 \sigma_x^3 \rangle = -1 \tag{7.7}$$

$$E(y, x, y) = \langle \sigma_y^1 \sigma_x^2 \sigma_y^3 \rangle = +1 \tag{7.8}$$

$$E(y, y, x) = \langle \sigma_y^1 \sigma_y^2 \sigma_x^3 \rangle = +1 \tag{7.9}$$

where σ_x^n and σ_y^n are the spin operators of particle n for spin along the x and y axes respectively. These are perfect correlations and, therefore, for each λ we have

$$A_1(x, \lambda) A_2(y, \lambda) A_3(y, \lambda) = +1 \tag{7.10}$$

$$A_1(x, \lambda) A_2(x, \lambda) A_3(x, \lambda) = -1 \tag{7.11}$$

$$A_1(y, \lambda) A_2(x, \lambda) A_3(y, \lambda) = +1 \tag{7.12}$$

$$A_1(y, \lambda) A_2(y, \lambda) A_3(x, \lambda) = +1 \tag{7.13}$$

Now, taking the product of the equations (7.10) to (7.13) gives a positive quantity on the LHS and a negative quantity on the RHS. This contradiction demonstrates that quantum mechanics is nonlocal without using inequalities. This demonstration does not involve any inequalities and it applies to every run of the experiment. The comparison with the statements $s1$ to sK of section 4.4 is clear indicating how that argument was motivated.

7.3 N-particle Bell inequalities by the CHSH method

In a real experiment the detectors would not be 100% efficient and therefore we would not have the perfect correlation between measurements of equations (7.6) to (7.9). Furthermore, the measurement apparatus would not be ideal and so that sometimes a + result would be recorded when a – result should have been recorded. These factors would block the GHZ demonstration when applied to a real experiment. Therefore the GHZ demonstration of nonlocality can only be regarded as applying to ideal experiments. To demonstrate nonlocality in real experiments inequalities would, once again, be required. In addition to being required in real experiments, inequalities would further serve to elucidate the relationship between the GHZ type demonstration and the Bell inequality type of demonstration of quantum nonlocality.

In this section we will derive N-particle inequalities by the CHSH method reviewed in section 4.3. To do this we put

$$E(a) - E(a') = \int \left(\prod_{n=1}^{N} A_n(a_n, \lambda) - \prod_{n=1}^{N} A_n(a'_n, \lambda) \right) \rho(\lambda) d\lambda$$

$$\begin{aligned} &= \int \left(\prod_{n=1}^{N} A_n(a_n, \lambda) \left[1 \pm \prod_{n=1}^{N} A_n(a'''_n, \lambda) \right] \right) \rho(\lambda) d\lambda \\ &\quad - \int \left(\prod_{n=1}^{N} A_n(a'_n, \lambda) \left[1 \pm \prod_{n=1}^{N} A_n(a''_n, \lambda) \right] \right) \rho(\lambda) d\lambda \end{aligned} \tag{7.14}$$

where we have chosen the a''_n and a'''_n so that,

$$A_n(a_n, \lambda) A_n(a'''_n, \lambda) = A_n(a'_n, \lambda) A_n(a''_n, \lambda) \tag{7.15}$$

for all n (see below). From (7.1) we see that the quantities in square brackets in (7.15) are positive. Therefore, using (7.1) and (7.4), we obtain

$$|E_(a) - E(a')| \leq 2 \pm (E(a''') + E(a'')) \tag{7.16}$$

Rearranging gives

$$-2 \leq E(a) - E(a') + E(a'') + E(a''') \leq 2 \tag{7.17}$$

Inequalities (7.17) together with equation (7.15) are N-measurement Bell inequalities. Equation (7.15) defines the possible relationships between a_n, a'_n, a''_n and a'''_n. Using

(7.15) we find that, for each value of n, there are the following possibilities:

$$a'_n = a_n \qquad \text{and} \qquad a'''_n = a''_n \tag{7.18}$$

or

$$a''_n = a_n \qquad \text{and} \qquad a'''_n = a'_n \tag{7.19}$$

In the case when there are no null results (7.2) holds so that we have the additional possibility,

$$a'''_n = a_n \qquad \text{and} \qquad a''_n = a'_n \tag{7.20}$$

To obtain the 2-particle Bell inequalities from the N-particle inequalities we put $n = 1, 2$. From (7.18) we can put,

$$a = a_1 = a'_1 \qquad \text{and} \qquad a' = a''_1 = a'''_1$$

and from (7.19) we can put,

$$b = a_2 = a''_2 \qquad \text{and} \qquad b' = a'_2 = a'''_2$$

Substituting these into (7.17) gives

$$-2 \leq E(a,b) - E(a,b') + E(a',b) + E(a',b') \leq 2 \tag{7.21}$$

These are the CHSH Bell inequalities derived in chapter 4.

To derive the 2-particle inequalities we only had to use possibilities (7.2) and (7.19). As we will see later, to derive useful three particle inequalities we will need to use possibility (7.20). However, this requires that there are no null results. To remove this requirement we will use a simple mathematical trick. Consider the measurement $A'_n(a_n)$ for which

$$\begin{aligned} A'_n(a_n,\lambda) =& A_n(a_n,\lambda) \quad \text{if} \quad A_n(a_n,\lambda) = \pm 1 \\ A'_n(a_n,\lambda) =& B_n(\lambda) \qquad\quad \text{if} \quad A_n(a_n,\lambda) = 0 \end{aligned} \tag{7.22}$$

where $B_n(\lambda)$ is generated randomly to take the values $+1$ and -1 each with 50% probability. For this measurement there are no null results. It is clear that the derivation of inequalities (7.17) in section 2 still holds when we substitute $A_n(a_n,\lambda)$ with $A'_n(a_n,\lambda)$

where, now, the correlation coefficient is

$$E'(a) = \int \Big(\prod_{n=1}^{N} A'_n(a_n, \lambda)\Big)\rho(\lambda)d\lambda \tag{7.23}$$

But, from (7.4) and (7.22)

$$E'(a) = E(a) + \int \Omega(M(\lambda))\Big(\prod_{n\in M(\lambda)} B_n(\lambda) \prod_{n\notin M(\lambda)} A_n(a_n, \lambda)\Big) \tag{7.24}$$

where $M(\lambda)$ is the set of n for which $A_n(a_n, \lambda) = 0$ and $\Omega(M(\lambda))$ is equal to 0 when the set $M(\lambda)$ is empty and equal to 1 when this set is not empty. However, the second term on the RHS of (7.24) will vanish because $B_n(\lambda)$ is generated randomly. Hence, $E'(a) = E(a)$ and therefore, from the point of view of the correlation function, making measurements of $A_n(a_n)$ (with null results) is equivalent to making measurements of $A'_n(a_n)$ (without null results). Therefore, inequalities obtained from (7.17) using (7.20) apply to experiments where there are null results and, consequently, can be used in real experiments.

7.4 N-particle Bell inequalities by a new method

In this section we will see how the method of section 4.4 can be used to obtain N-particle Bell inequalities. Consider K statements, s$'$1 to s$'K$, of the form

$$\text{s}'k \qquad \prod_{n=1}^{N} A_n(a_n^k, \lambda) = \begin{Bmatrix} \pm 1 & \text{if} & 1 \le k \le k_0 \\ \mp 1 & \text{if} & k_0 < k \le K \end{Bmatrix}$$

where k_0 is odd and K is even. Note that with $N = 2$, $k_0 = K - 1$, and appropriately chosen a_n^k 's these statements are the same as the sk statements of section 4.4. If we choose the a_n^k 's so that, for each n,

$$\prod_{k=1}^{K} A_n(a_n^k, \lambda) = 1 \quad \text{or} \quad 0 \tag{7.25}$$

(see below) then we have a contradiction because the product of all the statements s$'$1 to s$'K$ gives a +1 or 0 on the LHS (using (7.1) and -1 on the RHS (because k_0 is odd). Let $P^{\pm}$ be the probability that all the statements s$'$1 to s$'K$ are true and let $p_k^{\pm}(a^k)$ (shorthand for $p_k^{\pm}(a_1^k, a_2^k, ...a_N^k))$ be the probability that

$$\prod_{n=1}^{N} A_n(a_n^k, \lambda) = \pm 1$$

Now, the probability that one or more of the statements are false is less than or equal to the sum of the probabilities for each individual statement being false. That is,

$$1 - P^{\pm} \leq \sum_{k=1}^{k_0}(1 - p_k^{\pm}(a^k)) + \sum_{k=k_0+1}^{K} (1 - p_k^{\mp}(a^k)) \tag{7.26}$$

This simplifies to

$$P^{\pm} \geq \sum_{k=1}^{k_0} p_k^{\pm}(a^k) + \sum_{k=k_0+1}^{K} p_k^{\mp}(a^k) - (K-1) \tag{7.27}$$

If $P > 0$ then all of the statements s$'1$ to s$'K$ Proceeding as before, we obtain the more general Bell inequality

$$\sum_{k=1}^{K_0} p_k^{\pm}(a^k) + \sum_{k=k_0+1}^{K} p_k^{\mp}(a^k) \leq K - 1 \tag{7.28}$$

where k_0 is odd, K is even, and the a_n^k 's are chosen so that the condition (7.25) is satisfied. The condition (7.25) will be satisfied if the a_n^k 's can be grouped in pairs such that, for each n, there exist a set of integers k_{n1}, k_{n2}, . . . k_{nK} for which it is possible to write

$$a_n^{k_{n1}} = a_n^{k_{n2}}, \quad a_n^{k_{n3}} = a_n^{k_{n4}}, \quad . \quad . \quad . \quad a_n^{n_{k(K-1)}} = a_n^{k_{nK}} \tag{7.29}$$

where $k_{nj} \neq k_{nl}$ if $j \neq l$.

We will now show that it is possible to obtain CHSH type inequalities by using the method used to obtain inequalities (7.28) First we consider an experiment with no null results for which we will have $p_k^+ + p_k^- = 1$, or expressed differently,

$$p_k^{\pm} = 1 - p_k^{\mp} \tag{7.30}$$

Substituting this into inequality (7.28) we obtain

$$-\left(\sum_{k=1}^{k_0} p_k^{\mp}(a^k) + \sum_{k=k_0+1}^{K} p_k^{\pm}(a^k)\right) \leq -1 \tag{7.31}$$

Adding (7.28) and (7.31) gives

$$2 - K \leq \sum_{k=1}^{k_0} E(a^k) - \sum_{k=k_0+1}^{K} E(a^k) \leq K - 2 \tag{7.32}$$

where

$$E(a^k) = p_k^+(a^k) - p_k^-(a^k) \tag{7.33}$$

If $K = 4$ and $k_0 = 3$ then (7.32) is the CHSH type inequality (7.17). This derivation works equally well backwards, that is we could obtain (7.28) from (7.32) using (7.20). At this stage we have only considered experiments in which there are no null results. The requirement that there are no null results can be removed by the same trick we used in the previous section. Therefore inequalities (7.32) apply as they are written even when there are null results.

7.5 Connection between GHZ and inequalities

We can construct 3-particle inequalities that apply to the GHZ demonstration given in section 7.2 by putting, using (7.18) to (7.20)

$$a_1 = a_1' = x \qquad \text{and} \qquad a_1'' = a_1''' = y$$

$$a_2' = a_2'' = x \qquad \text{and} \qquad a_2 = a_2''' = y$$

$$a_3' = a_3''' = x \qquad \text{and} \qquad a_3 = a_3'' = y$$

in inequalities (7.17) giving

$$-2 \leq E(x, y, y) - E(x, x, x) + E(y, x, y) + E(y, y, x) \leq 2 \tag{7.34}$$

The upper bound of (7.34) can also be obtained by putting $N = 3$ in Mermin's inequalities (1990a) (see also the inequalities of Roy and Singh (1991)). From (7.6) to (7.9) we see that

$$E(x, y, y) - E(x, x, x) + E(y, x, y) + E(y, y, x) = 4$$

Since $|E| \leq 1$, this constitutes a maximum possible violation of inequality (7.34). In this case, the maximum violation of the Bell inequalities corresponds to the GHZ contradiction between quantum mechanics and locality without inequalities. In fact, in those circumstances where a maximum violation of the Bell inequalities can be obtained, such a contradiction will always exist. To prove this we will consider inequalities (7.32) which are more general than inequalities (7.17). First we note that we get a maximum violation of inequalities (7.32) if, and only if,

$$E(a^{k'}) = -E(a^{k''}) = \pm 1 \tag{7.35}$$

for $k' = 1$ to k_0 and $k'' = k_0 + 1$ to K. This gives, using (7.4)

$$E(a^{k'}) = \prod_{n=1}^{N} A_n(a_n^{k'}, \lambda) = \pm 1 \tag{7.36}$$

$$E(a^{k''}) = \prod_{n=1}^{N} A_n(a_n^{k''}, \lambda) = \mp 1 \tag{7.37}$$

It is not necessary to sum over λ when there is a perfect correlation. These equations require that there are no null results and consequently equation (7.25) becomes

$$\prod_{k=1}^{K} A_n(a_n^k, \lambda) = 1 \tag{7.38}$$

for $k = 1$ to K. Taking the product of equations (7.36) for all k' and (7.37) for all k'' gives

$$\prod_{n=1}^{N} \prod_{k=1}^{K} A_n(a_n^k, \lambda) = -1$$

However, by substituting in (7.38) we find that the LHS is positive but that the RHS is negative. This is a GHZ type contradiction.

7.6 Preparing N-atom entangled states with single photons

The method of preparing a two-atom entangled state using a Mach-Zehnder interferometer with a single photon source was discussed in the previous chapter. Here we will show that is possible to extend this approach to prepare a N-atom entangled state using $N - 1$ interferometers and a single photon source. These entangled states will be used to illustrate the violation of the N-particle Bell inequalities by the predictions of quantum mechanics.

Recall from section 6.3 that spin $\frac{1}{2}$ atoms are prepared in the initial state

$$|\psi_n\rangle = \beta_{n+}|+\rangle_n + \beta_{n-}|-\rangle_n \tag{7.39}$$

They are then each are placed in a box, a nonuniform magnetic field is placed along the boxes (which are orientated along the z-axis) and a wall is placed so as to divide each box into two and then the nonuniform magnetic field is removed so that the state of each atom plus box system becomes

$$|\Psi_n\rangle = \beta_{n+}|A,+\rangle_n + \beta_{n-}|B,-\rangle_n \tag{7.40}$$

We will have N of these atom plus box systems, i.e. $n = 1$ to N. Furthermore we will have a number of Mach-Zehnder interferometers, MZI, $MZII$, etc each prepared so that if a photon is fed into the input then due to destructive interference it will not emerge out of output d_m (the 'dark' output)for the m interferometer when the paths through it are unobstructed.

To prepare a two atom entangled state one atom plus box system is placed in each of the two paths of the interferometer such that the path goes through the A_n box (see fig. 6.2). When the photon is detected in the d_1 output it was shown in section 6.3 that the (unnormalized) state of the two atoms becomes

$$|\Phi_{12}\rangle = \beta_{1+}\beta_{2-}|+\rangle_1|-\rangle_2 - \beta_{1-}\beta 2+|-\rangle_1|+\rangle_2 \tag{7.41}$$

where we have now removed the dividing walls of the boxes. If the atoms are initially prepared so that they have spin $+\frac{1}{2}$ along the x-axis then the initial state of each atom is

$$|\psi_n\rangle = \frac{1}{\sqrt{2}}(|+\rangle_n + |-\rangle_n) \tag{7.42}$$

Therefore, $\beta_{n+} = \beta_{n-} = \frac{1}{\sqrt{2}}$. Hence, when the photon is detected in the d_1 output the normalized sate vector of the atoms is, from (7.41),

$$|\Phi_{12}\rangle = \frac{1}{\sqrt{2}}(|+\rangle_1|-\rangle_2 - |-\rangle_1|+\rangle_2) \tag{7.43}$$

This is a singlet state. Suitable measurements on this state will violate the Bell inequalities.

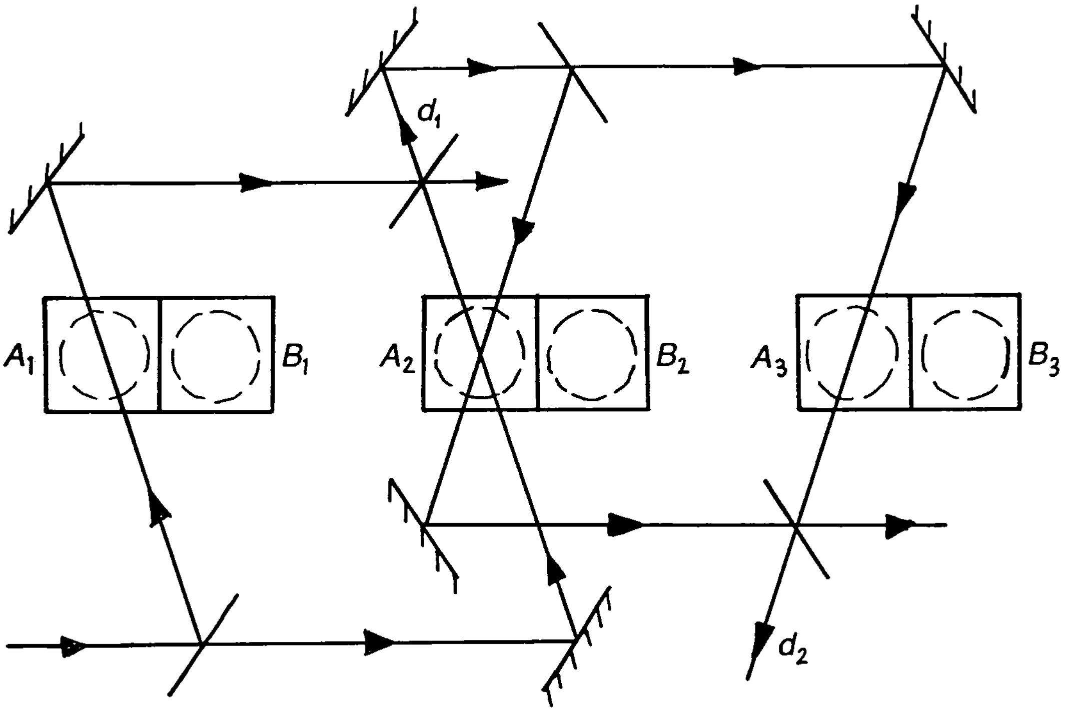

Fig. 7.1. Two entangled Mach-Zehnder interferometers arranged to produce a three atom entangled state.

To produce a three particle entangled state we can use the arrangement shown in fig. 7.1. This arrangement consists of two Mach-Zehnder interferometers placed so that the 'dark' output of the first, MZI, goes into the input of the second, MZII. Furthermore, one arm of each interferometer goes through box A_2 of atom 2 so that MZI has the atoms 1 and 2, one in each arm, and MZII has the atoms 2 and 3, one in each arm. We shall refer to two interferometers when arranged in this way as *entangled.*

If the photon emerges at output d_1, then from (7.40) the unnormalized state vector of the atoms is

$$|\Phi_{12}\rangle = (-\beta_{1-}\beta_{2+}|-\rangle_1)|+\rangle_2 + (\beta_{1+}\beta_{2-}|+\rangle_1)|-\rangle_2 \tag{7.44}$$

The photon now goes into MZII. The state of atom 2 is now given by (7.44) where the quantities in brackets are to be regarded as coefficients. The state of atom 3 is

$$|\psi_3\rangle = \beta_{3+}|+\rangle_3 + \beta_{3-}|-\rangle_3, \tag{7.45}$$

If the photon now emerges at the 'dark' output d_2 then the states of atoms 2 and 3 will become entangled. As the states of atoms 1 and 2 are already entangled we find that we have the 3 atom entangled state,

$$|\Phi_{123}\rangle = -\beta_{1+}\beta_{2-}\beta_{3+}|+\rangle_1|-\rangle_2|+\rangle_3 - \beta_{1-}\beta_{2+}\beta_{3-}|-\rangle_1|+\rangle_2|-\rangle_3 \tag{7.46}$$

We obtain (7.46) from (7.41) by making the changes

$$|\pm\rangle_1 \longrightarrow |\pm\rangle_2$$

$$|\pm\rangle_2 \longrightarrow |\pm\rangle_3$$

and, comparing (7.44) and (7.45) with (7.39)

$$\beta_{1\pm} \longrightarrow \mp\beta_{1\mp}\beta_{2\pm}|\mp\rangle_1$$

$$\beta_{2\pm} \longrightarrow \beta_{3\pm}$$

The photon could now be fed into a third interferometer, MZIII, entangled with MZII such that atom 3 is in an arm of both MZII and MZIII. We would then have a 4-atom entangled state. In principle, we can have as many entangled interferometers as we wish. Preparing all the atoms in the initial state (7.42) so that $\beta_{n\pm} = \frac{1}{\sqrt{2}}$, we get the normalised N-atom state,

$$|\varphi\rangle_N = \frac{1}{\sqrt{2}}(|+-+-...\rangle_N - (-1)^N|-+-+...\rangle_N \tag{7.47}$$

by using $N-1$ entangled interferometers. In the notation used in (7.47), $\pm$ in the nth position corresponds to a spin of $\pm\frac{1}{2}$ along the z-axis for the nth atom. Consider, now, the correlation function

$$E(a) = {}_N\langle\varphi|\left(\prod_{n=1}^{N}\sigma_n(a_n)\right)|\varphi\rangle_N \tag{7.48}$$

where $\sigma_n(a_n)$ is the spin operator corresponding to a measurement of spin on atom n along a direction in the xy-plane at an angle a_n to the x-axis. Using elementary quantum mechanics, we can show that

$$E(a) = (-1)^{N-1}\cos(a_1 - a_2 + a_3 - ...a_N) \tag{7.49}$$

It is well established that the Bell inequalities are violated by the predictions of quantum theory for $N = 2$. We will now show that, according to quantum theory, appropriate measurements on the state (7.42) will give a maximum violation of the N-measurement Bell inequalities for all $N \geq 3$.

Using (7.18) to (7.20) we put,

$$a_1 = a_1' = 0 \qquad \text{and} \qquad a_1'' = a_1''' = \frac{\pi}{2}$$

$$a_2' = a_2'' = 0 \qquad \text{and} \qquad a_2 = a_2''' = -\frac{\pi}{2}$$

$$a_3' = a_3''' = 0 \qquad \text{and} \qquad a_3 = a_3'' = \frac{\pi}{2}$$

and, for all $k > 3$

$$a_n = a_n' = a_n'' = a_n''' = 0$$

Using (7.17) we obtain the Bell inequalities

$$\begin{aligned} -2 \leq E(0, -\frac{\pi}{2}, \frac{\pi}{2}, 0, 0, ...0) - E(0, 0, 0, 0, 0, ...0) \\ +E(\frac{\pi}{2}, 0, \frac{\pi}{2}, 0, 0, ...0) + E(\frac{\pi}{2}, -\frac{\pi}{2}, 0, 0, 0, ...0) \leq 2 \end{aligned} \tag{7.50}$$

From (7.48) we get,

$$E(o, -\frac{\pi}{2}, \frac{\pi}{2}, 0, 0, ...0) = (-1)^N \tag{7.51}$$

$$E(0, 0, 0, 0, 0, ...0) = -(-1)^N \tag{7.52}$$

$$E(\frac{\pi}{2}, 0, \frac{\pi}{2}, 0, 0, ...0) = (-1)^N \tag{7.53}$$

$$E(\frac{\pi}{2}, -\frac{\pi}{2}, 0, 0, 0, ...0) = (-1)^N \tag{7.54}$$

These give a maximum violation of the Bell inequalities (7.50). Therefore, we can have a maximum violation of the Bell inequalities for all $N \geq 3$. The GHZ type contradiction can be demonstrated whenever there is a maximum violation of Bell's inequalities and, therefore, can be demonstrated, by using these N-atom states, for all $N \geq 3$. In fact it is only the first three atoms that play an important role in producing the violation of the Bell inequalities here. This is a fault of the inequalities. It has been shown by Pagonis, Redhead and Clifton (1991) that it is possible to produce a GHZ contradiction for any number of particles with each particle playing an important role. Furthermore, inequalities have been derived by Mermin (1991b) and by Roy and Singh (1991) which do exploit each particle even when $N > 3$. The principle value of the N-particle inequalities derived in this chapter is that they illustrate how the methods of chapter 4 can be extended from 2 particles to N particles.

7.7 Preparing N-atom entangled states by shuffling boxes

In section 5.4 we saw how a two atom entangled state (or strictly speaking, a two ion entangled state) can be prepared by 'shuffling boxes'. In this section we will see that the same method can be extended to prepare N-atom entangled states.

We take N ions, $n = 1$ to N. The ions for which n is odd are positively charged and the ions for which n is even are negatively charged. When a positive and a negative ion are in the same box they combine irreversibly to produce a molecule. We will consider, first the case $N = 3$. As with the case $N = 2$ discussed in section 5.4, each ion is prepared in the state $|\psi_n\rangle$ for $n = 1$ to 3 (see equation (7.39)above).

Next each ion is placed into a box (see fig. 7.2(a)) and a nonuniform magnetic field is applied, dividing walls are placed so as to divide each box into two boxes and then the magnetic field is removed (fig. 7.2(b)) so that the state of each ion plus boxes system becomes $|\Psi_n\rangle$ (see equation (7.40) above). The state of the system at this stage is

$$\prod_{n=1}^{3} |\Psi_n\rangle \tag{7.55}$$

Next, boxes A_1 and A_2 are brought together and the walls that divide them are removed and likewise boxes B_1 and B_2 are brought together and the walls that divide them are removed (see fig. 7.2(c)). If ion 1 and ion 2 are in the same box then they will combine to form a molecule. As before, we assume that the walls of the box are transparent to the neutral molecules and consequently any molecule produced will fall out of the box under gravity. Now an electric field is applied across each box such that ion 1 which is positive goes back into box A_1 or B_1 and ion 2 which is negative goes back into box A_2 or B_2. The dividing walls can now be replaced and then the electric field removed (see fig. 7.2(d)). This leaves the system in the unnormalized state

$$\Big(\beta_{1+}\beta_{2-}|A,+\rangle_1|B,-\rangle_2 - \beta_{1-}\beta 2{+}|B,-\rangle_1|A,+\rangle_2\Big)|\Psi_3\rangle \tag{7.56}$$

where we have projected out the terms with a molecule. Now, this process is repeated with ions 2 and 3 as shown in figs. 7.2(e) to 7.2(g). This will leave the system in the unnormalized state $|\Phi_{123}\rangle$ in equation (7.46). Thus we can prepare the same entangled state as in the the previous section. Also, as before, we can extend this to N ions by repeating the 'box shuffle' described above with ions 3 and 4 and then ions 4 and 5 etc. until the desired number of ions is reached.

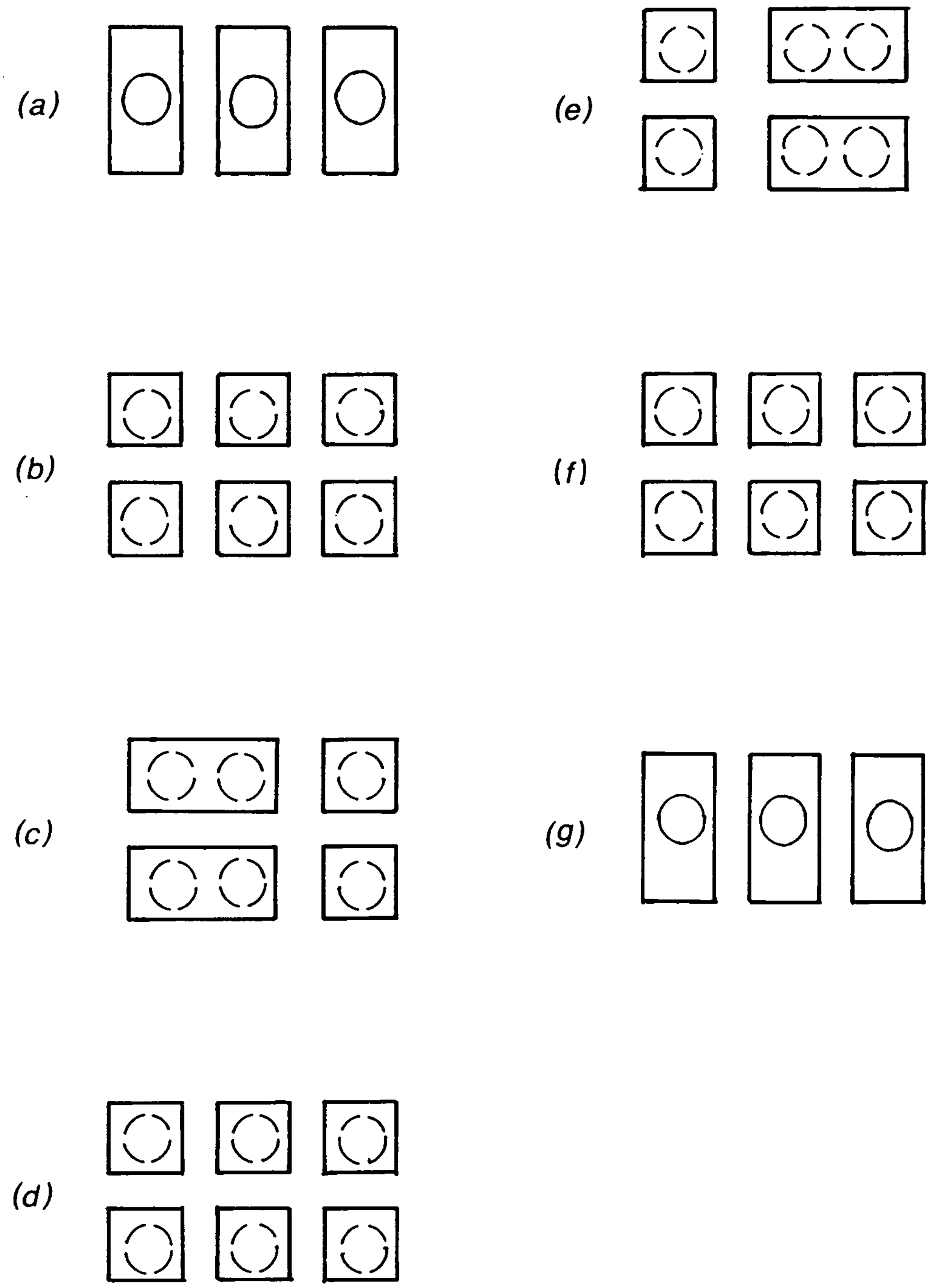

Fig. 7.2. The procedure for preparing a three atom entangled state by shuffling boxes.

Chapter 8

Bell's Theorem Without Inequalities for Two Particles

8.1 Introduction

The inequality-free demonstration of quantum nonlocality of GHZ discussed in the previous chapter required three or more particles. Except in the case where the number of settings of the local variable is allowed to tend to infinity as discussed in chapter 4, it is not possible to use the GHZ approach to demonstrate quantum nonlocality with only two particles. Heywood and Redhead (1967) have demonstrated nonlocality for two spin one particles without using inequalities by using a Bell-Kochen-Specker type argument. Both the GHZ method and the Heywood and Redhead method require a total of six Hilbert space dimensions. In this chapter we will show how another approach does allow us to demonstrate Bell's theorem without using inequalities with only two particles each particle living in a two dimensional Hilbert space so that we have a total of only four dimensions in our Hilbert space. However, unlike the demonstration of GHZ where nonlocality was demonstrated to be a property of every run of the experiment, in this case nonlocality is only demonstrated in $\frac{1}{16}$th of runs of the experiment.

To run this demonstration of Bell's theorem we will consider a gedanken experiment which is a modification of the gedanken experiment considered in chapter 3 in the discussion of empty waves. This new gedanken experiment consists of two overlapping Mach-Zehnder-type interferometers. We will then show how the same arguments can instead be run in the context of a 'shuffling boxes' experiment. Neither of these proposals could be realized in the laboratory. Thus we will also consider a quantum optical version of the experiment that could be realized. However, in an actual experiment the conditions necessary to run the demonstration without inequalities will no longer apply. Therefore, it will be necessary to use inequalities. The necessary inequalities are derived.

8.2 Nonlocality with two overlapping interferometers

Recall that in the gedanken experiment considered in chapter 3 in connection with empty waves, an atom plus box system is placed with one of its boxes in one path of a Mach-Zehnder interferometer. The path lengths inside the interferometer are set so that one output is dark if there are no obstructions in the internal paths. The atom is initially prepared with spin equal to $+\frac{1}{2}$ along the x-axis. If the atom plus box system is not placed with one box in the path of the interferometer then when the two boxes are brought back together and a measurement of spin along the x-axis is made then the value $+\frac{1}{2}$ will definitely be recorded. The atom plus box system can be seen to be playing the same role as the interferometer. Thus, a result of $-\frac{1}{2}$ is analogous to a detection in the dark output. Given this, we can replace the atom plus box system by another interferometer. This way the apparatus is made to look more symmetrical. Thus, although we could use the apparatus of chapter 3, we will instead use the apparatus shown in fig. 8.1. This consists of two Mach-Zehnder-type interferometers MZ_i ($i = 1, 2$), one for positrons (MZ_1) and one for electrons (MZ_2), arranged so that two paths overlap as shown in fig. 8.1. Each interferometer MZ_i has an input mode, s_i, two paths inside the interferometer, u_i and v_i, two output modes, c_i and d_i and two beam splitters BSA_i and BSB_i. Taken separately each interferometer is arranged so that, due to destructive interference, no positrons/electrons will be detected at detector D_i in output d_i. The beam splitters BSB_i are removable. Now, a positron and an electron are created simultaneously and fed into their respective interferometers. The apparatus is arranged such that, if the positron takes path u_1 inside MZ_1 and the electron takes path u_2 inside MZ_2, then the two particles will meet at point P and annihilate one another with probability equal to one. Expressing this mathematically we have,

$$|u_1\rangle|u_2\rangle \longrightarrow |\gamma\rangle, \tag{8.1}$$

where $|u_i\rangle$ is the state of the positron/electron travelling along path u_i and $|\gamma\rangle$ is the state of the radiation produced on annihilation. We will find that, as a consequence of this possible interaction between the two particles, it becomes possible for positrons/electrons to arrive at detectors D_i.

The operation of BSA_i is given by,

$$|s_i\rangle \longrightarrow \frac{1}{\sqrt{2}}\Big(i|u_i\rangle + |v_i\rangle\Big) \tag{8.2}$$

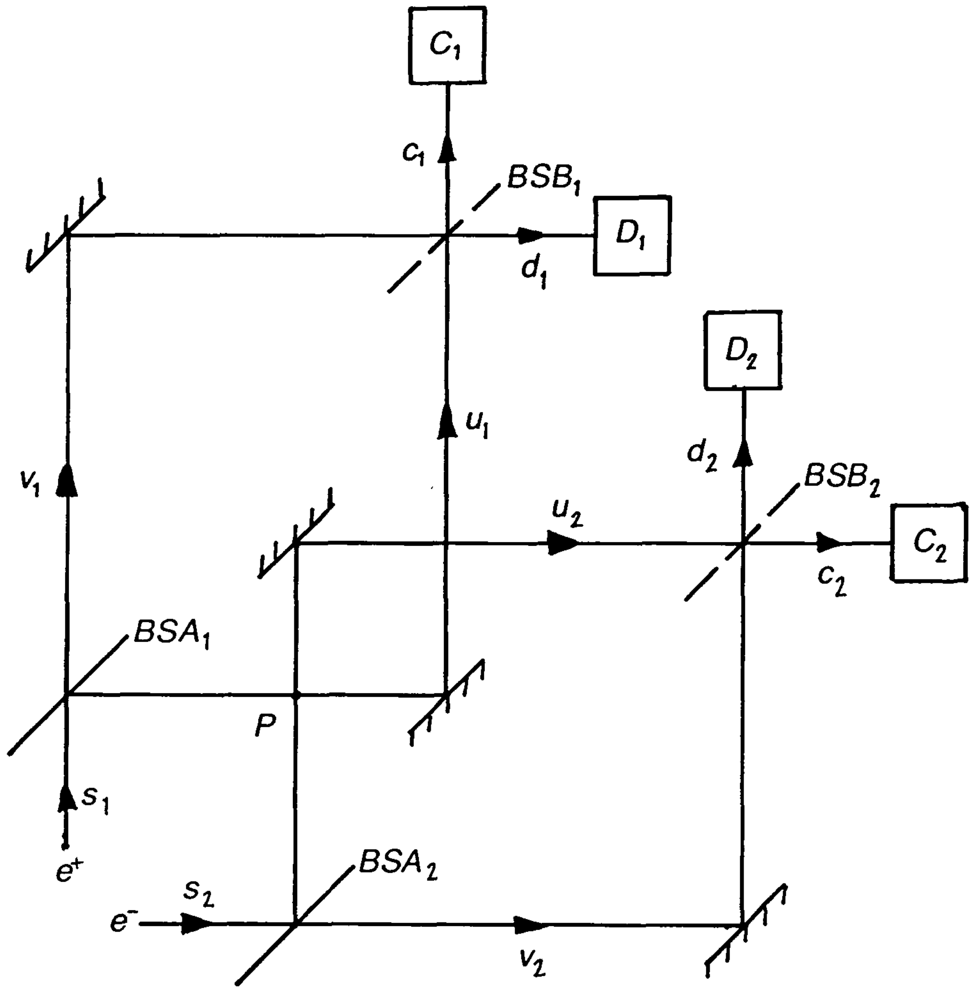

Fig. 8.1. Two Mach-Zehnder-type interferometers, one for positrons and one for electrons, arranged such that if a positron takes path u_1 and an electron takes path u_2 then they will meet at point P and annihilate one another.

The operation of BSB_i is given by,

$$|u_i\rangle \longrightarrow \frac{1}{\sqrt{2}}\Big(|c_i\rangle + i|d_i\rangle\Big) \tag{8.3}$$

and,

$$|v_i\rangle \longrightarrow \frac{1}{\sqrt{2}}\Big(i|c_i\rangle + |d_i\rangle\Big) \tag{8.4}$$

If BSB_i is removed then,

$$|u_i\rangle \longrightarrow |c_i\rangle \tag{8.5}$$

and

$$|v_i\rangle \longrightarrow |d_i\rangle \tag{8.6}$$

The initial state of the system is,

$$|s_1\rangle|s_2\rangle \tag{8.7}$$

After passing through the beam splitters BSA_i this state evolves to (using (8.2)),

$$\frac{1}{2}\Big(i|u_1\rangle + |v_1\rangle\Big)\Big(i|u_2\rangle + |v_2\rangle\Big) \tag{8.8}$$

After passing point P the state becomes, using (8.1),

$$\frac{1}{2}\Big(- |\gamma\rangle + i|u_1\rangle|v_2\rangle + i|v_1\rangle|u_2\rangle + |v_1\rangle|v_2\rangle\Big) \tag{8.9}$$

If both BSB_1 and BSB_2 removed then, using (8.5) and (8.6), we see that (8.9) evolves to the final state,

$$\frac{1}{2}\Big(- |\gamma\rangle + i|c_1\rangle|d_2\rangle + i|d_1\rangle|c_2\rangle + |d_1\rangle|d_2\rangle\Big) \tag{8.10}$$

With BSB_1 in place and BSB_2 removed, using (8.3) to (8.6) we find that (8.9) evolves to the final state,

$$\frac{1}{2\sqrt{2}}\Big(- \sqrt{2}|\gamma\rangle - |c_1\rangle|c_2\rangle + 2i|c_1\rangle|d_2\rangle + i|d_1\rangle|c_2\rangle\Big) \tag{8.11}$$

Similarly, with BSB_1 removed and BSB_2 in place we find that (8.9) evolves to the final state,

$$\frac{1}{2\sqrt{2}}\Big(- \sqrt{2}|\gamma\rangle - |c_1\rangle|c_2\rangle + i|c_1\rangle|d_2\rangle + 2i|d_1\rangle|c_2\rangle\Big) \tag{8.12}$$

If both beam splitters BSB_i are in place then using (8.3) and (8.4) we find that (8.9) evolves to the final state,

$$\frac{1}{4}\Big(- 2|\gamma\rangle - 3|c_1\rangle|c_2\rangle + i|c_1\rangle|d_2\rangle + i|d_1\rangle|c_2\rangle - |d_1\rangle|d_2\rangle\Big) \tag{8.13}$$

The notion of realism is introduced by letting the state of the positron-electron pair before measurements are made be described by hidden variables, λ. These hidden variables can take different values each time the experiment is repeated. We can one of make two measurements on each particle – either with the beam splitter in place, denoted by 0, or with the beam splitter removed, denoted by ∞. The assumption of locality requires that the result of a measurement on one particle does not depend on the choice of measurement on the other particle. If the positron/electron is detected at D_i with beam splitter BSB_i in place then we will write $D_i(0,\lambda) = 1$, if it is not detected then we will write $D_i(0,\lambda) = 0$. If the positron/electron is detected at C_i with the beam splitter BSB_i removed then we will write $C_i(\infty,\lambda) = 1$, if it is not detected then we will write $C_i(\infty,\lambda) = 0$. In adopting this notation we have assumed locality because the result of a measurement on one particle does not depend on the choice of measurement made on the other particle. For example, $D_1(0,\lambda)$ does not depend on whether BSB_2 is in place or not. We will now see that this leads to a contradiction with quantum mechanics. From (8.10) we see that,

$$C_1(\infty,\lambda)C_2(\infty,\lambda) = 0 \tag{8.14}$$

for every experiment because there is no $|c_1\rangle|c_2\rangle$ term. From (8.11) we see that,

$$\text{if} \qquad D_1(0,\lambda) = 1 \qquad \text{then} \qquad C_2(\infty,\lambda) = 1 \tag{8.15}$$

because if the positron is detected at D_1 then the state is projected on to the last term in (8.11). Similarly, from (8.12) we see that,

$$\text{if} \qquad D_2(0,\lambda) = 1 \qquad \text{then} \qquad C_1(\infty,\lambda) = 1 \tag{8.16}$$

From (8.13) we see that

$$D_1(0,\lambda)D_2(0,\lambda) = 1 \qquad \text{for } \frac{1}{16}\text{th of experiments.} \tag{8.17}$$

Now consider an experiment for which $D_1(0,\lambda)D_2(0,\lambda) = 1$ From (8.17) we see that this will happen in $\frac{1}{16}$th of experiments. From (8.15) and (8.16) we see this implies that $C_1(\infty,\lambda)C_2(\infty,\lambda) = 1$ for these experiments. However, (8.14) tells us that $C_1(\infty,\lambda)C_2(\infty,\lambda) = 0$ for all experiments. Hence we have a contradiction between local realism and quantum mechanics. Whilst this result can be compared to the GHZ result because no inequalities are used it is dissimilar in that it only applies to $\frac{1}{16}$th of the experiments whereas the GHZ result applies to every run of the experiment.

8.3 Inferring the existence of the result functions

In a comment on the above proof, Greenberger it et. al. (GBHZ) have pointed out that the existence of the result functions $C_i(0,\lambda)$ and $D_i(\infty,\lambda)$ are assumed rather than being proven because there is not perfect correlation between certain measurements at the two ends. In the case of the singlet state considered by Bell and in the case of the three particle state considered by GHZ one can deduce what the result of a measurement of some spin component will be by making appropriate measurements on the other particle(s). Take the example of a singlet state. Bell states "Since we can predict in advance the result of measuring any chosen component of σ_2, by previously measuring the same component of σ_1, it follows that the result of any such measurement must actually be predetermined" because the assumption of locality dictates that "the orientation of one magnet does not influence the result obtained with the other." Therefore, it is not necessary to assume the existence of the functions $A(\mathbf{a},\lambda)$ (where this is the value that would be obtained if σ_1 was measured along the **a** direction) etc. By contrast, in the demonstration in the previous section there is not perfect correlation between the values that are measured and hence GHBZ claim that it is not possible to prove that functions $C_1(\infty,\lambda)$, etc. exist by rerunning Bell's argument.

It should first be noted that this is not new to Bell theory. For example, the correlation function $E(a,b)=\frac{9}{10}\cos(a-b)$ violates the CHSH inequalities and therefore is regarded as violating local realism even though there is not a perfect correlation between the measurements at the two ends. The point to be made here is that even when we cannot prove the existence of the result functions it is still impossible to explain any correlation that violates the CHSH inequalities by a local hidden variable interpretation. Indeed the very possibility of performing experiments depends on the fact that the inequalities can be used when there is not perfect correlation between the measurements at the two ends.

However, as we shall now see, it is possible to prove the existence of $C_i(\infty,\lambda)$ in the appropriate subensemble and thereby run the nonlocality argument without having to make the assumption about existence of functions like $C_1(\infty,\lambda)$ that GHBZ have identified in the above proof. We will not need to prove or assume the existence of the result functions $D_i(0,\lambda)$. We will consider the actual context of the experiment to be that in which the beam splitters are in place and we will consider a run of the experiment for which we obtain measurement results $D_1=1$ and $D_2=1$. Now, as $D_1=1$, we can predict that a measurement of $C_2(\infty)$ would certainly yield the value 1 from (8.11). Furthermore, that it has this value cannot depend on the fact that we have chosen to put the beam splitter BSB_1 in place because we assume locality. Therefore, by EPR type reasoning, we must have $C_2(\infty, \lambda) = 1$ for this run of the experiment. We have, then, proved the existence of the result function $C_1(\infty, \lambda)$. One point should be made clear here. We do not assume that if, contrary to fact, we had decided to perform the experiment with BSB_1 in place and BSB_2 removed we would have still certainly obtained $D_1 = 1$ for this run of the experiment. Such an assumption cannot be deduced from locality except in the case of determinism (see page 92-3 of Redhead (1987) for discussion of this). For our purposes it is enough that we might have obtained $D_1 = 1$. Therefore the electron end of the apparatus cannot 'know' that we will certainly not get $D_1 = 1$ for this run of the experiment and consequently we must have $C_1(\infty, \lambda) = 1$ so that there is no possibility of the predictions of quantum mechanics being violated. We can use the same reasoning to establish that for this run of the experiment we have $C_1(\infty, \lambda) = 1$. Therefore, for this run of the experiment we have established that $C_1(\infty, \lambda)C_2(\infty, \lambda) = 1$ but this contradicts (8.14). The reason that it is possible to run this argument is because we have *result dependent perfect correlation.*

In one sense, this nonlocality result is stronger than previous results because it is only necessary to prove the existence of two result functions but it is weaker in that these result functions cannot always be proven to exist. However, its greatest strength is its simplicity: it does not require inequalities (the GHZ demonstration also does not require inequalities) and it only requires two particles (GHZ require three particles).

8.4 Illustrating nonlocality with trajectories

The above demonstration of nonlocality can be illustrated in a very striking way by assuming the existence of particle trajectories. In fact we will only assume the existence of trajectories in those situations for which we can infer what the actual paths taken by the particles are. Consider some hidden variables for which both $D_1(0, \lambda) = 1$ and $D_2(0, \lambda) = 1$. Suppose that the detection at D_1 happens before the electron reaches BSB_2. Now making a measurement of $C_2(\infty)$ (i.e with beam splitter BSB_2 removed) is equivalent to placing a detector into the u_2 path. However, from (8.15) we see that, for our choice of hidden variables, if a detector was to be placed in the u_2 path then it would detect the electron. Therefore, we will assume that, even if the detector is not placed in the path, the electron is actually goes along the u_2 path. If the electron takes the u_2 path then the positron must have taken the v_1 path otherwise it would have met the electron at point P and annihilated and could not then have been detected at detector D_1. Therefore the actual paths taken when both beam splitters are in place are those shown in fig. 8.2(a).

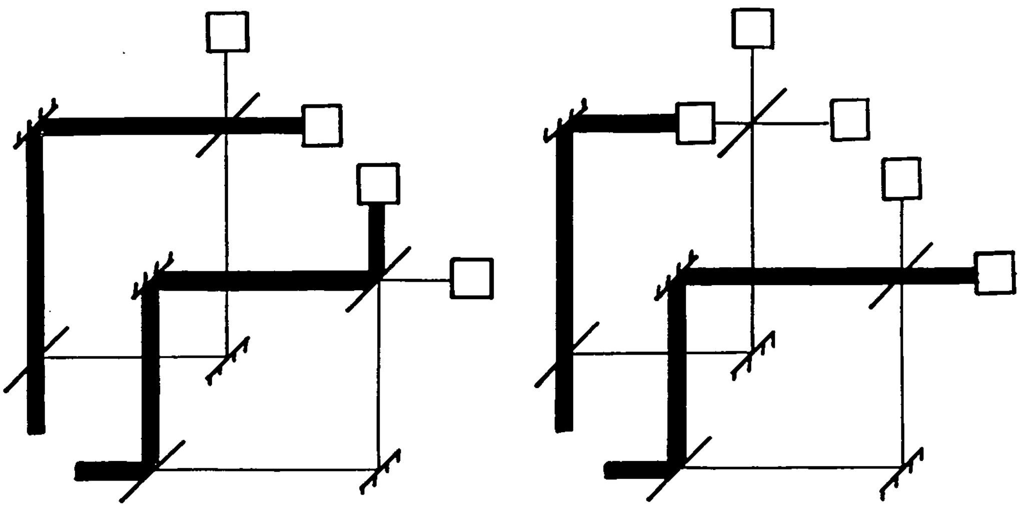

Fig. 8.2. Illustration of nonlocality with trajectories.

The arguments used here are essentially the same as the arguments used in chapter 2 with respect to the apparatus considered there. Now consider what would happen if a detector was placed into the v_1 path to intercept the positron. The positron would be detected at this detector. However, if the positron is detected in the v_1 path then this effectively decouples the two interferometers and consequently the electron cannot now be detected in the dark output. Thus for the same λ as in fig. 8.2(a) we now have the trajectories shown in fig. 8.2(b). We see quite clearly that the path taken by the electron depends on what is measured on the positron even though the choice of what is measured on the positron can be made in a region spacelike separated from the positron.

8.5 Shuffling boxes method

Instead of using interferometers to run the nonlocality argument above one can use the 'shuffling boxes' method first discussed in section 5.4. and also discussed in the previous chapter. The particular shuffle required this time is shown in fig. 8.3. We take two ions, 1 and 2. Ion 1 is positively charged and ion two is negatively charged. If the two ions are at any time in the same box then they will combine irreversibly to form a molecule. The ions are prepared so that they have spin $+\frac{1}{2}$ along the x-axis. Thus their initial states are

$$|\psi_n\rangle = \frac{1}{\sqrt{2}}(|+\rangle_n + |-\rangle_n) \tag{8.18}$$

for $n = 1, 2$. They are then each placed in a box, a nonuniform magnetic field is placed along the length of each box which is orientated along the z-axis, dividing walls are put in place and the nonuniform magnetic field is removed. This puts each of the ion plus boxes systems into the state

$$|\Psi_n\rangle = \frac{1}{\sqrt{2}}(|A,+\rangle_n + |B,-\rangle_n) \tag{8.19}$$

At this stage the two systems have not interacted and thus the state of the whole system is

$$\frac{1}{2}\Big(|A,+\rangle_1 + |B,-\rangle_1\Big)\Big(|A,+\rangle_2 + |B,-\rangle_2\Big) \tag{8.20}$$

(see fig. 8.3(b)). Now box A_1 is brought together with box A_2 and the dividing walls removed to form box A. If the ions are both in this box then they will combine to form a molecule, that is

$$|A,+\rangle_1|A,+\rangle_2 \longrightarrow |\text{molecule}\rangle \tag{8.21}$$

As previously we will assume that the walls of the box are transparent to the neutral molecule and consequently, the molecule will drop out of the box under gravity. A electric field is put across box A such that if ion 1 is in the box then it will go to the A_1 side of the box and if ion two is in the box then it will go to the A_2 side. Next the dividing walls are replaced and the electric field removed. The state of the system will now have evolved from that in (8.20) to

$$\frac{1}{2}\Big(|\text{molecule}\rangle + |A,+\rangle_1|B,-\rangle_2 + |B,-\rangle_1|A,+\rangle_2 + |B,-\rangle_1|B,-\rangle_2\Big) \tag{8.22}$$

where we have used (8.21). This state is similar to the state in (8.9). Ion one and ion two are separated so that measurements can be made on them at a spacelike distance.

One of two measurements can be made on each ion plus boxes system. We can look to see if the ion is in box A_n or B_n. This is effectively a measurement of spin along the z-axis. Alternatively, we can bring the box A_n and the box B_n together, remove the dividing walls and then make a measurement of spin along the x axis. We have the result function $P_{n\pm}(z,\lambda)$ which is equal to 1 if ion n has spin $\pm\frac{1}{2}$ along the z-axis and equal to 0 otherwise. Similarly, we have the result function $P_{n\pm}(x,\lambda)$ which is equal to 1 if ion n has spin $\pm\frac{1}{2}$ along the x-axis and equal to 0 otherwise. We will write the state in (8.20) as

$$\frac{1}{2}\Big(|\text{molecule}\rangle + |+\rangle_1|-\rangle_2 + |-\rangle_1|+\rangle_2 + |-\rangle_1|-\rangle_2\Big) \tag{8.23}$$

where we have suppressed the A's and B's as they are redundant. The z-axis spin eigen-

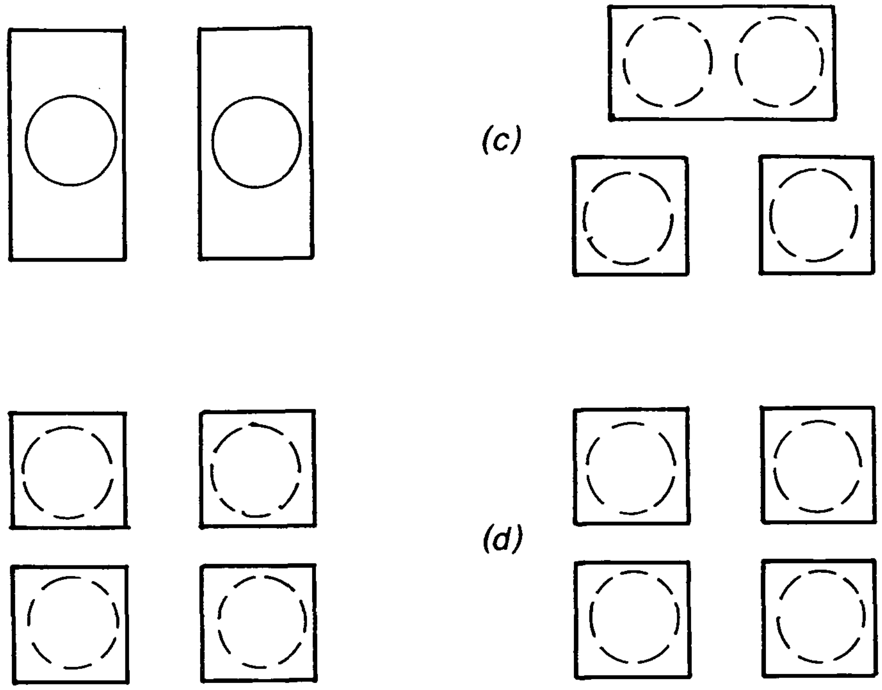

Fig. 8.3. The box shuffle for producing the two-atom entangled state that can be used to demonstrate Bell's theorem without inequalities.

states can be written in terms of the x-axis spin eigenstates,

$$|\pm\rangle_n = \frac{1}{\sqrt{2}}\Big(|+\rangle_{nx} \pm |-\rangle_{nx}\Big) \qquad (8.24)$$

Consider a measurement of spin along the z-axis on both ions (this is analogous to removing both beam splitters in the interferometer example above). From (8.23) we see immediately that

$$P_{1+}(z,\lambda)P_{2+}(z,\lambda) = 0 \qquad (8.25)$$

as there is no $|+\rangle_1|+\rangle_2$ term in (8.23) (cf. equation (8.14)). To see what happens if x spin is measured on ion 1 and z spin is measured on ion 2 we can rewrite (8.23) as

$$\frac{1}{2\sqrt{2}}\Big(\sqrt{2}|\text{molecule}\rangle + |+\rangle_{1x}|+\rangle_2 + 2|+\rangle_{1x}|-\rangle_2 - |-\rangle_{1x}|+\rangle_2 \qquad (8.26)$$

where we have used equation (8.24). From (8.26) we see that

$$\text{if} \qquad P_{1-}(x,\lambda) = 1 \qquad \text{then} \qquad P_{2+}(z,\lambda) = 1, \qquad (8.27)$$

(cf. equation (8.15)). If x spin is measured on ion 2 and z spin is measured on ion 1 then by symmetry we see from (8.27) that

$$\text{if} \qquad P_{2-}(x,\lambda)=1 \qquad \text{then} \qquad P_{1+}(z,\lambda)=1, \tag{8.28}$$

(cf. equation (8.16)). Finally to see what happens if x spin is measured on both ions we rewrite (8.23) as

$$\frac{1}{4}\Big(2|\text{molecule}\rangle + 3|+\rangle_{1x}|+\rangle_{2x} - |+\rangle_{1x}|-\rangle_{2x} - |-\rangle_{1x}|+\rangle_{2x} - |-\rangle_{1x}|-\rangle_{2x}\Big) \tag{8.29}$$

From (8.29) we see that

$$P_{1-}(x,\lambda)P_{2-}(x,\lambda)=1 \qquad \text{for } \frac{1}{16}\text{th of experiments} \tag{8.30}$$

(cf. equation (8.17)). The demonstration of nonlocality can now be run as before. Thus, we consider an experiment for which $P_{1-}(x,\lambda)P_{2-}(x,\lambda) = 1$. From (8.30) we see that this will happen sometimes. From (8.27) and (8.28) we see that this implies that $P_{1+}(z,\lambda)P_{2+}(z,\lambda) = 1$ but this contradicts equation (8.25). Therefore, we have a contradiction between quantum theory and local realism.

This gedanken experiment is essentially the spin analogue of the interferometer experiment although the boxes play an important role in the preparation of the state (8.23). The spin analogue of this experiment has also been considered by Clifton and Niemann (1992) who generalise the demonstration to two spin s particles and by Pagonis and Clifton (1992) who generalise the demonstration to N spin $\frac{1}{2}$ particles.

8.6 A quantum optical version

We would not expect it to be possible to realise either the overlapping interferometers version or the shuffling boxes version of this nonlocality demonstration in a real experiment. Thus, we will consider a quantum optical version that could be realised. However, in a real experiment we will not see the perfect correlations of gedanken experiments. Therefore it is necessary to derive inequalities that could be applied in a real experiment. In the limit of an ideal experiment these inequalities become redundant. Furthermore, it is necessary to make supplementary assumptions analogous to those discussed in section 5.2 because present technology cannot provide us with sufficiently high efficiency detectors.

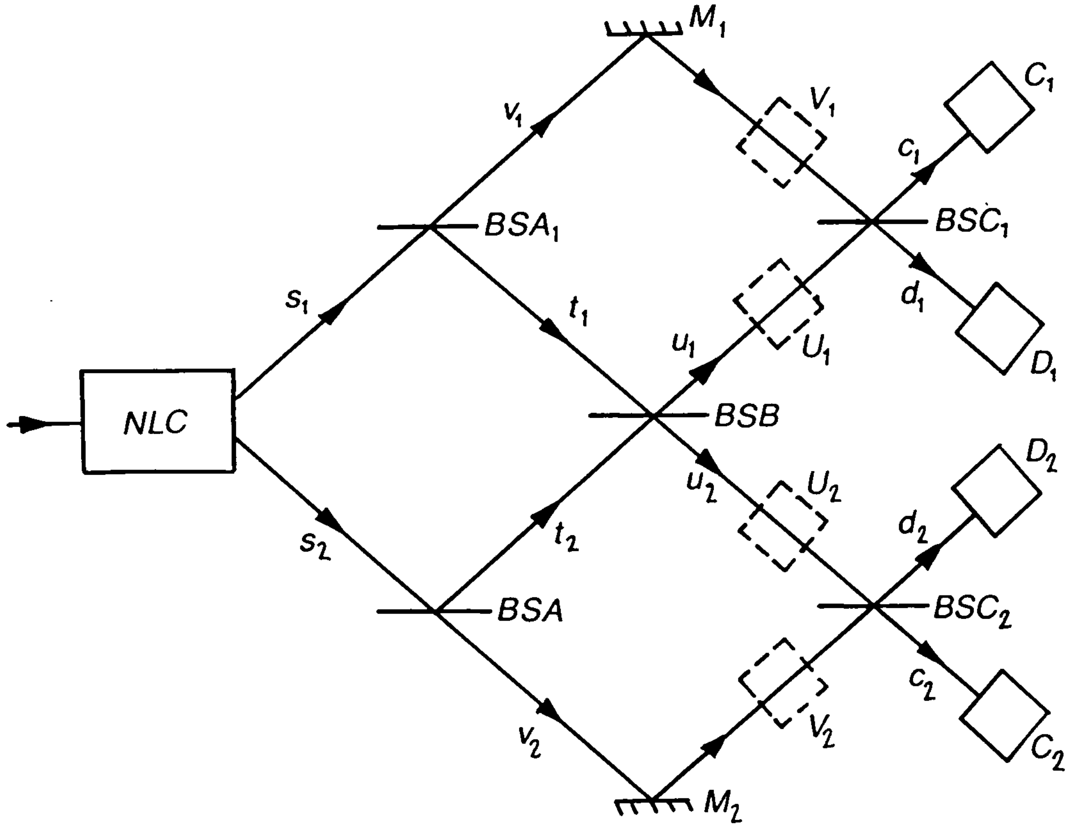

Fig. 8.4. Photon pairs are created by degenerate parametric down conversion. Each photon of the pair is fed into an interferometer system. The two interferometer systems meet at the beam splitter BSB.

The proposed experiment is shown in fig. 8.4. Pairs of photons with the same frequency are produced by degenerate parametric down conversion at the nonlinear crystal NLC. Each photon, 1 and 2, of each pair is fed into an interferometer system. The interferometer system for photon i ($i = 1, 2$) consists of the beam splitter BSA_i, the mirror M_i, and the beam splitter BSC_i. These two interferometer systems meet in one of their corners at beam splitter BSB which has transmittance equal to $\frac{1}{2}$. The first beam splitter, BSA_i, of each interferometer system has transmittance equal to $\frac{1}{2}$. The last beam splitter, BSC_i, of each interferometer system has transmittance equal to $\frac{1}{3}$. With these choices for the transmittance of the beam splitters it is possible to set the path lengths such that, if photon two (one) is blocked from entering its interferometer, then photon one (two) cannot be detected at the D_1 (D_2) detector because of destructive interference. In the following we will assume that the path lengths have been set in this way. We will see that the effect of the interaction of the two photons at BSB is to make it possible for photons to be detected at both D_1 and D_2. Instead of performing this interference

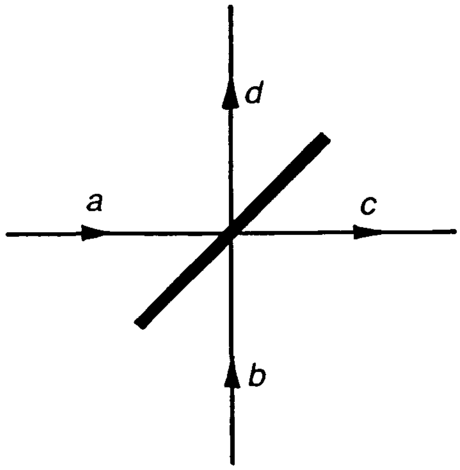

Fig. 8.5. A beam splitter with inputs a and b and outputs c and d.

experiment, it is possible instead to insert detectors U_i and V_i into the internal paths of the interferometers as shown in fig. 8.4. Thus, at each end of the apparatus there are two experiments that can be performed: Either the interference experiment which we will denote by $e_i = d_i$ or the experiment with U_i and V_i inserted which we will denote by $e_i = u_i$. If n photons are detected at detector D_i then we put $D_i = n$, and similarly for the other detectors. We are assuming that the detectors are capable of distinguishing between one photon and two photons. Define $N_i(e_i)$ as the total number of photons detected at end i when experiment e_i is performed. Thus

$$N_i(d_i) = C_i + D_i \quad \text{and} \quad N_i(u_i) = U_i + V_i \tag{8.31}$$

First we will derive some basic results for photons impinging on beam splitters. If $\hat{a}$ is the annihilation operator for mode a then we have the result

$$|n\rangle_a = \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}}|0\rangle \tag{8.32}$$

Fig. 8.5. Shows a beam splitter with input modes a and b and output modes c and d. The input and output annihilation operators are related by

$$\begin{pmatrix} \hat{c} \\ \hat{d} \end{pmatrix} = \begin{pmatrix} \sqrt{T} & i\sqrt{R} \\ i\sqrt{R} & \sqrt{T} \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} \tag{8.33}$$

with inverse

$$\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} \sqrt{T} & -i\sqrt{R} \\ -i\sqrt{R} & \sqrt{T} \end{pmatrix} \begin{pmatrix} \hat{c} \\ \hat{d} \end{pmatrix} \tag{8.34}$$

where R and T are the reflectance and transmittance respectively (see Campos, Saleh and

Teich (1989)). Consider a single photon in the a mode. The state of the system is

$$|1\rangle_a|0\rangle_b = a^\dagger|0\rangle_a|0\rangle_b$$

This evolves to, using (8.32) and (8.34),

$$(\sqrt{T}\hat{c}^\dagger + i\sqrt{R}\hat{d}^\dagger)|0\rangle_c|0\rangle_d = \sqrt{T}|1\rangle_c|0\rangle_d + i\sqrt{R}|0\rangle_c|1\rangle_d$$

That is

$$|1\rangle_a|0\rangle_b \longrightarrow \sqrt{T}|1\rangle_c|0\rangle_d + i\sqrt{R}|0\rangle_c|1\rangle_d \tag{8.35}$$

Now consider what happens when there is a single photon in each of the input modes a and b with $R = T = \frac{1}{2}$. The input state is

$$|1\rangle_a|1\rangle_b = \hat{a}^\dagger\hat{b}^\dagger|0\rangle_a|0\rangle_b$$

This evolves to, using (8.32), (8.34), and the fact that $\hat{c}^\dagger$ commutes with $\hat{d}^\dagger$ for photons,

$$\frac{1}{2}(\hat{c}^\dagger + i\hat{d}^\dagger)(i\hat{c}^\dagger + \hat{d}^\dagger)|0\rangle_c|0\rangle_d = \frac{i}{\sqrt{2}}\Big(|2\rangle_c|0\rangle_d + |0\rangle_c|2\rangle_d\Big)$$

That is

$$|1\rangle_a|1\rangle_b \longrightarrow \frac{i}{\sqrt{2}}\Big(|2\rangle_c|0\rangle_d + |0\rangle_c|2\rangle_d\Big) \tag{8.36}$$

This shows that, with a single photon incident on each input of a 50:50 beam splitter, both photons must emerge out of the same output. It is not possible for one to emerge at the c output and the other to emerge at the d output. This effect has been experimentally verified by Hong, Ou, and Mandel (1987). Finally, at mirrors M_i we assume a phase change of $\frac{\pi}{2}$, that is

$$|1\rangle_{v_i} \longrightarrow i|1\rangle_{v_i} \tag{8.37}$$

Note that this is not an essential assumption. If the phase change at the mirrors is not $\frac{\pi}{2}$ then a phase shifter could be placed in path v_i so that the combined effect of the mirror and phase shifter is to bring about an effective phase shift of $\frac{\pi}{2}$. This is simply to ensure that, in the calculation below, the phase difference between the two paths through each of the two interferometer systems is equal to zero or a multiple of 2π so that there is destructive interference in output d_1 (d_2) when input s_2 (s_1) is blocked.

We will now apply these results to the apparatus in fig. 8.4. The initial state of the system is

$$|1\rangle_{s_1}|1\rangle_{s_2} \tag{8.38}$$

After passing through the beam splitters BSA_i the state becomes, using (8.35),

$$\frac{1}{2}\left(i|1\rangle_{t_1}|0\rangle_{v_1} + |0\rangle_{t_1}|1\rangle_{v_1}\right)\left(i|1\rangle_{t_2}|0\rangle_{v_2} + |0\rangle_{t_2}|1\rangle_{v_2}\right) \tag{8.39}$$

On passing through BSB and reflecting off mirrors M_i the state becomes, using (8.35) to (8.37),

$$\frac{-1}{2\sqrt{2}}\left(|2\rangle_{u_1}|0\rangle_{u_2} + |0\rangle_{u_1}|2\rangle_{u_2}\right)|0\rangle_{v_1}|0\rangle_{v_2} - \frac{1}{2\sqrt{2}}\left(i|1\rangle_{u_1}|0\rangle_{u_2} + |0\rangle_{u_1}|1\rangle_{u_2}\right)|0\rangle_{v_1}|1\rangle_{v_2}$$

$$-\frac{1}{2\sqrt{2}}\left(|1\rangle_{u_1}|0\rangle_{u_2} + i|0\rangle_{u_1}|1\rangle_{u_2}\right)|1\rangle_{v_1}|0\rangle_{v_2} - \frac{1}{2}|0\rangle_{u_1}|0\rangle_{u_2}|1\rangle_{v_1}|1\rangle_{v_2}$$

Simplifying gives

$$\frac{1}{\sqrt{2}}|\varphi\rangle - \frac{i}{2\sqrt{2}}|1\rangle_{u_1}|0\rangle_{u_2}|0\rangle_{v_1}|1\rangle_{v_2} - \frac{i}{2\sqrt{2}}|0\rangle_{u_1}|1\rangle_{u_2}|1\rangle_{v_1}|0\rangle_{v_2} - \frac{1}{2}|0\rangle_{u_1}|0\rangle_{u_2}|1\rangle_{v_1}|1\rangle_{v_2} \tag{8.40}$$

where

$$|\varphi\rangle = \frac{-1}{2}\Big(|2\rangle_{u_1}|0\rangle_{u_2}|0\rangle_{v_1}|0\rangle_{v_2} + |0\rangle_{u_1}|2\rangle_{u_2}|0\rangle_{v_1}|0\rangle_{v_2}$$

$$+|0\rangle_{u_1}|1\rangle_{u_2}|0\rangle_{v_1}|1\rangle_{v_2} + |1\rangle_{u_1}|0\rangle_{u_2}|1\rangle_{v_1}|0\rangle_{v_2}\Big) \tag{8.41}$$

The state $|\varphi\rangle$ only contains terms with both photons in the same interferometer. We shall only be interested in those runs of the experiment for which one photon goes to end 1 and the other goes to end 2, i.e. for which $N_1N_2 = 1$. Therefore, we need not pay any special attention to the evolution of the state $|\varphi\rangle$. From expression (8.40) we see that if the detectors U_1, V_1, U_2, and V_2 are inserted then

$$U_1U_2 = 0 \tag{8.42}$$

for all runs of the experiment as there is no $|1\rangle_{u_1}|1\rangle_{u_2}|0\rangle_{v_1}|0\rangle_{v_2}$ term. If the detectors U_2 and V_2 are inserted but the detectors U_1 and V_1 are not, then the state in (8.40) evolves through BSC_1 but not through BSC_2 giving, using (8.35),

$$\frac{1}{\sqrt{2}}|\varphi_1\rangle - \frac{i\sqrt{3}}{2\sqrt{2}}|1\rangle_{c_1}|0\rangle_{d_1}|0\rangle_{u_2}|1\rangle_{v_2} + \frac{1}{2\sqrt{3}}|1\rangle_{c_1}|0\rangle_{d_1}|1\rangle_{u_2}|0\rangle_{v_2} - \frac{i}{2\sqrt{6}}|0\rangle_{c_1}|1\rangle_{d_1}|1\rangle_{u_2}|0\rangle_{v_2} \tag{8.43}$$

where $|\varphi_1\rangle$ is the state resulting from evolving $|\varphi\rangle$ through BSC_1 but not BSC_2. Only the last term in (8.43) contains $|0\rangle_{c_1}|1\rangle_{d_1}$. Therefore

$$U_2 = 1 \quad \text{if} \quad D_1 = 1 \quad \text{and} \quad N_2(u_2) = 1 \tag{8.44}$$

If the detectors $U_1$and V_1 are inserted but the detectors U_2 and V_2 are not, then by symmetry, we have

$$U_1 = 1 \quad \text{if} \quad D_2 = 1 \quad \text{and} \quad N_1(u_1) = 1 \tag{8.45}$$

Finally, if detectors U_1, V_1, U_2, and V_2 are all removed, then the state in (8.40) evolves through both BSC_1 and BSC_2 to give

$$\frac{1}{\sqrt{2}}|\varphi_{12}\rangle + \frac{2}{3}|1\rangle_{c_1}|0\rangle_{d_1}|1\rangle_{c_2}|0\rangle_{d_2} - \frac{i}{6\sqrt{2}}|1\rangle_{c_1}|0\rangle_{d_1}|0\rangle_{c_2}|1\rangle_{d_2}$$

$$-\frac{i}{6\sqrt{2}}|0\rangle_{c_1}|1\rangle_{d_1}|1\rangle_{c_2}|0\rangle_{d_2} + \frac{1}{6}|0\rangle_{c_1}|1\rangle_{d_1}|0\rangle_{c_2}|1\rangle_{d_2} \tag{8.46}$$

where $|\varphi_{12}\rangle$ is the state resulting from evolving $|\varphi\rangle$ through both BSC_1 and BSC_2. From (8.46) state we see that

$$D_1 D_2 = 1 \quad \text{for} \quad \frac{1}{18}\text{th} \quad \text{of} \quad \text{experiments} \quad \text{in} \quad \text{which} \quad N_1(d_1)N_2(d_2) = 1 \tag{8.47}$$

We will now show how this experiment can be used to demonstrate that realistic interpretations of quantum mechanics must be nonlocal. First we will consider the ideal case. The notion of realism is introduced by letting the state of the photon pair before the measurements are made be described by the set of hidden variables λ. These hidden variables determine the results of measurements that are made. Locality requires that the result of a measurement at one end does not depend on the choice of measurement at the other end. Consequently, the results of performing experiment d_i at end i are $D_i(\lambda)$ and

$C_i(\lambda)$ with

$$N_i(d_i, \lambda) = C_i(\lambda) + D_i(\lambda)$$

Similarly, the results of performing experiment u_i at end i are $U_i(\lambda)$ and $V_i(\lambda)$ with

$$N_i(u_i, \lambda) = U_i(\lambda) + V_i(\lambda)$$

In adopting this notation we have assumed locality. For example, $D_1(\lambda)$ does not depend on whether experiment d_2 or experiment u_2 is performed at end 2. For each run of the experiment there are two photons. Therefore,

$$N_1(e_1, \lambda) + N_2(e_2, \lambda) = 2 \tag{8.48}$$

If e_j is held fixed and e_i $(i \neq j)$ is varied then the l.h.s. of (8.48) must remain constant and equal to the r.h.s. Thus we have

$$N_i(d_i, \lambda) = N_i(u_i, \lambda) \tag{8.49}$$

From (8.42), (8.44), (8.45), and (8.47) we have

$$U_1(\lambda)U_2(\lambda) = 0 \tag{8.50}$$

$$U_2(\lambda) = 1 \quad \text{if} \quad D_1(\lambda) = 1 \quad \text{and} \quad N_2(u_2, \lambda) = 1 \tag{8.51}$$

$$U_1(\lambda) = 1 \quad \text{if} \quad D_2(\lambda) = 1 \quad \text{and} \quad N_1(u_1, \lambda) = 1 \tag{8.52}$$

$$D_1(\lambda)D_2(\lambda) = 1 \quad \text{for} \quad \frac{1}{18}\text{th} \quad \text{of} \quad \text{expmts.} \quad \text{in} \quad \text{which} \quad N_1(d_1, \lambda)N_2(d_2, \lambda) = 1 \tag{8.53}$$

The results (8.49) to (8.53) give rise to a contradiction. Consider a run of the experiment for which $D_1(\lambda)D_2(\lambda) = 1$. From (8.53) we see that this will happen sometimes. If $D_1(\lambda)D_2(\lambda) = 1$ then $N_i(d_i, \lambda) = 1$ for $i = 1$ and 2. Therefore, from (8.49) we have $N_i(u_i, \lambda) = 1$ also for $i = 1$ and 2. From (8.51) and (8.52) we see that $D_1(\lambda)D_2(\lambda) = 1$ and $N_i(u_i, \lambda) = 1$ imply that $U_1(\lambda)U_2(\lambda) = 1$ but this contradicts (8.50). Hence, local realistic interpretations of quantum mechanics are not possible.

We will now show how inequalities can be obtained that can be applied in real experiments. First consider an experiment with ideal detectors but in which the rest of the apparatus is not ideal such that the predictions (8.42), (8.44), (8.45), and (8.47) do not hold exactly. This could be because the process indicated in (8.36) does not hold exactly. In the experiment of Hong, Ou, and Mandel (1987) it held in approximately 90% of cases.

Another source of non-ideality could be that the phase difference is not exactly set to ensure full destructive interference in output d_1 (d_2) when input s_2 (s_1) is blocked. This can be compared with the results of Grangier, Roger, and Aspect (1986) who achieved 98% visibility in a Mach-Zehnder interferometer with a single photon source. As we have ideal detectors, we will still have a total of two photons detected for each run of the experiment. Therefore, equations (8.48) and (8.49) will still hold.

We consider statements of the form

$$S_{AB} \qquad\qquad A_1(\lambda)B_2(\lambda) = 1$$

where $A, B = D, U, N$. These statements are either true (if $A_1(\lambda)B_2(\lambda) = 1$) or false. Let $\sim S_{AB}$ be the statement that S_{AB} is false, that is

$$\sim S_{AB} \qquad\qquad A_1(\lambda)B_2(\lambda) = 0$$

Notice that statements involving N implicitly use equation (8.49). For example, S_{ND} is the statement that $N_1(\lambda)D_2(\lambda) = 1$ but, here, N_1 does not depend on whether $e_1 = d_1$ or u_1.

Let Q, R, T, and W be statements each of which can be identified with statements of the form S_{AB}. We will consider probabilities like $P(Q\&R/T\&W)$ where this is the probability that Q and R are true given that T and W are true. More precisely, this probability is equal to the number of runs of the experiment for which λ is such that Q, R, T and W are true divided by the number of runs for which λ is such that T and W is true, in the limit as the number of runs tends to infinity. We will now derive two inequalities that will be required later. The statement R is either true or false, therefore,

$$P(Q/T) = P(Q\&R/T) + P(Q\& \sim R/T) \tag{8.54}$$

but

$$P(Q\& \sim R/T) \leq P(\sim R/T)$$

which together with (8.54) gives

$$P(Q\&R/T) \geq P(Q/T) - P(\sim R/T) \tag{8.55}$$

Now,

$$P(Q/R\&T) = \frac{P(Q\&R/T)}{P(R/T)} \tag{8.56}$$

and

$$P(R/T) = 1 - P(\sim R/T) \tag{8.57}$$

Substituting (8.55) and (8.57) into (8.56) gives the inequality

$$P(Q/R\&T) \geq \frac{P(Q/T) - P(\sim R/T)}{1 - P(\sim R/T)} \tag{8.58}$$

This is the first of the inequalities that will be required later. To derive the second inequality we notice that

$$\begin{aligned} P(\sim \{Q\&R\&W\}/T) &\leq P(\sim Q\& \sim R\& \sim W/T) \\ &\leq P(\sim Q/T) + P(\sim R/T) + P(\sim W/T) \end{aligned}$$

Using (8.57) for R and similar relations for Q, W, and $\{Q\&R\&W\}$ gives

$$P(Q\&R\&W/T) \geq P(Q/T) + P(R/T) + P(W/T) - 2 \tag{8.59}$$

These results can now be applied to obtain a Bell inequality applicable to the quantum optical experiment considered above. If

$$P(S_{DD}) > 0 \tag{8.60}$$

and

$$P(\{\sim S_{UU}\}\&S_{DU}\&S_{UD}/S_{DD}) > 0 \tag{8.61}$$

then we have a contradiction. To see this, consider λ for which S_{DD}, $\sim S_{UU}$, S_{DU}, and S_{UD} are true. Such λ must exist if (8.60) and (8.61) hold. For such λ we have

$$D_1(\lambda)D_2(\lambda) = 1 \tag{8.62}$$

$$U_1(\lambda)U_2(\lambda) = 0 \tag{8.63}$$

$$D_1(\lambda)U_2(\lambda) = 1 \tag{8.64}$$

$$U_1(\lambda)D_2(\lambda) = 1 \tag{8.65}$$

The product of equations (8.62), (8.64), and (8.65) gives $U_1(\lambda)U_2(\lambda) = 1$ which contradicts equation (8.65). Using (8.59) we see that if

$$P(\sim S_{UU}/S_{DD}) + P(S_{DU}/S_{DD}) + P(S_{UD}/S_{DD}) > 2 \tag{8.66}$$

then (8.61) is satisfied. The probabilities in this inequality cannot be measured because they refer to incompatible measurement contexts. However, by using (8.58), we can place a lower limit on the values of these probabilities. For the first term in (8.66) we have

$$P(\sim S_{UU}/S_{DD}) = P(\sim S_{UU}/S_{DD}\&S_{NN}) \tag{8.67}$$

because S_{NN} is true if S_{DD} is true. Therefore, using (8.58) gives

$$P(\sim S_{UU}/S_{DD}) \geq \frac{P(\sim S_{UU}/S_{NN}) - P(\sim S_{DD}/S_{NN})}{1 - P(\sim S_{DD}/S_{NN})} \tag{8.68}$$

for the second term in (8.66) we have

$$P(S_{DU}/S_{DD}) = P(S_{DU}/S_{DD}\&S_{DN}) \tag{8.69}$$

because S_{DN} is true if S_{DD} is true. Using (8.58) gives

$$P(S_{DU}/S_{DD}) \geq \frac{P(S_{DU}/S_{DN}) - P(\sim S_{DD}/S_{DN})}{1 - P(\sim S_{DD}/S_{DN})} \tag{8.70}$$

Similarly, for the third term in (8.66) we have

$$P(S_{UD}/S_{DD}) \geq \frac{P(S_{UD}/S_{ND}) - P(\sim S_{DD}/S_{ND})}{1 - P(\sim S_{DD}/S_{ND})} \tag{8.71}$$

Using (8.68), (8.70), and (8.71) we see that (8.66) (and therefore (8.61)) is satisfied if

$$\begin{aligned}\frac{P(\sim S_{UU}/S_{NN}) - P(\sim S_{DD}/S_{NN})}{1 - P(\sim S_{DD}/S_{NN})} &+ \frac{P(S_{DU}/S_{DN}) - P(\sim S_{DD}/S_{DN})}{1 - P(\sim S_{DD}/S_{DN})} \\ &+ \frac{P(S_{UD}/S_{ND}) - P(\sim S_{DD}/S_{ND})}{1 - P(\sim S_{DD}/S_{ND})} > 2\end{aligned} \tag{8.72}$$

The probabilities in this inequality can be measured because they each refer to compatible measurements. If this inequality and inequality (8.60) above are satisfied then there is a contradiction with local realism. We see that the quantum mechanical predictions for an ideal experiment are, from (8.50) to (8.53),

$$P(\sim S_{UU}/S_{NN}) = P(S_{DU}/S_{DN}) = P(S_{UD}/S_{ND}) = 1$$

$$P(S_{DD}/S_{NN}) = \frac{1}{18}$$

and, from (8.46),

$$P(S_{DD}/S_{DN}) = P(S_{DD}/S_{ND}) = \frac{1}{12}$$

These values satisfy inequalities (8.60) and (8.72) reconfirming our earlier result that local realistic interpretations of quantum mechanics are not possible. If the apparatus is not ideal then the probabilities will not take these values but may still satisfy these inequalities. So far our discussion has only considered ideal detectors. If we do not have ideal detectors then we must make supplementary assumptions like those considered in chapter 5. We consider the 'true' values of the observables D_i and U_i to be the values that would be measured with ideal detectors. We then assume fair sampling such that, for events where one photon is detected at each end of the apparatus, the proportion of detections at detector D_i is the same as it would be with ideal detectors and similarly for the other detectors. If this is true then the measured values of the probabilities in (8.60) and (8.72) with non-ideal detectors are the same as the values that would be measured with ideal detectors. Thus the probabilities in the inequalities (8.60) and (8.72) can now be interpreted as the those that are measured in a real experiment. Without ideal detectors, it cannot be confirmed that the total number of photons per run of the experiment is equal to 2. Therefore, we must now regard equation (8.49) (which follows from equation (8.48)) as an supplementary assumption where N_i is the 'true' value for the number of photons at end i. Photomultipliers are usually used to detect photons in quantum optical experiments. These cannot distinguish between one photon and two photons. Fortunately, this does not matter for this proposal because we are only interested in those runs of the experiment in which one photon goes to each end. For these events, no more than one photon will impinge on any given detector.

When the inequalities (8.60) and (8.72) are satisfied, then local realism must be violated. However, with Bell inequalities, the contradiction is usually expressed the other way round. That is, if the Bell inequality is *violated* then local realism is violated. Such an inequality can be formed from (8.60) and (8.72):

$$\begin{aligned}\Bigg[&\frac{P(\sim S_{UU}/S_{NN}) - P(\sim S_{DD}/S_{NN})}{1 - P(\sim S_{DD}/S_{NN})} + \frac{P(S_{DU}/S_{DN}) - P(\sim S_{DD}/S_{DN})}{1 - P(\sim S_{DD}/S_{DN})} \\ &+ \frac{P(S_{UD}/S_{ND}) - P(\sim S_{DD}/S_{ND})}{1 - P(\sim S_{DD}/S_{ND})}\Bigg] P(S_{DD}) \leq 2P(S_{DD})\end{aligned} \tag{8.73}$$

If this inequality is violated then both (8.60) and (8.72) are satisfied, so that local realism is violated. Inequality (8.73) can be regarded as a Bell inequality.

The inequalities (8.60) and (8.72) (or equivalently (8.73)) could be tested in a real experiment. From estimates based on the results of the Hong, Ou, and Mandel experiment and also the experiment of Grangier, Roger, and Aspect, it would appear that such an experiment may be possible with present technology.

Whilst the main topic of this paper has been local realism, the particular quantum optical arrangement we have considered has another interesting property (besides exhibiting nonlocality) that deserves mention. If we only consider events in which one photon is detected at each end, then we know that the photon detected at end 1 has come from the source of photon 1, i.e. along path s_1 even though there also exists a path by which photon 2 could have reached end 1. This is because, for photon 2 to enter the interferometer of photon 1 it must be transmitted at beam splitter BSB. However, if this happens then photon 1 cannot at the same time enter into the interferometer of photon 2 because it would also have to be transmitted at BSB and we know from (8.36) that this cannot happen. Therefore, if photon 2 is in the interferometer of photon 1 then photon 1 must also be in this interferometer but we are considering the case where one photon is detected at each end. This is an unusual situation because when we have an experimental situation in which there are paths to a particular detector that are open to two or more identical particles it would usually be the case that, when a particle is detected at this detector, it is impossible to say which source this particular particle came from.

Another version of the two overlapping interferometers has been proposed by Yurke and Stoler (1992b) using fermions rather than bosons. In their apparatus the two identical fermions have a common path in the overlapping part of the interferometers. However, the Pauli exclusion principle forbids both fermions from taking this path serving the same function as the annihilation it section 8.2. However, fermion interferometers are difficult to construct and therefore it is unlikely that their proposal could be realized.

8.7 Conclusions

The new demonstration of Bell's theorem discussed in this chapter provides a particularly simple demonstration of nonlocality in quantum mechanics. We have considered two gedanken experiments that can be used in this demonstration and also one realizable quantum optical experiment.

Chapter 9

Lorentz Invariance in Quantum Theory

9.1 Introduction

The violation of locality discussed in previous chapters raises the question of whether there is an incompatibility between relativity and realistic interpretations of quantum mechanics. It may be thought that Bell's demonstration of the violation of locality in realistic interpretations of quantum theory is, in itself, sufficient to tell us that Lorentz invariance must also be violated in such interpretations. However, it is not at all clear that this must be the case. It has been argued by Ballentine and Jarrett (1987) that such an incompatibility cannot be demonstrated by using Bell's theorem. In this chapter we take a different approach. A condition for the existence of elements of reality and a condition for the Lorentz invariance of these elements of reality are defined. We will call interpretations for which both these conditions are true *Lorentz-invariant realistic Interpretations.* It is shown that when these two conditions are applied to quantum mechanics we obtain a contradiction.

We will consider the two overlapping interferometer gedanken experiment used in the previous chapter to demonstrate Bell's theorem for two particles without inequalities. The demonstration to be presented here has been generalized to two spin s particles by Clifton and Niemann (1992). A similar argument against Lorentz invariance is given by Pitowsky (1991) using the three spin $\frac{1}{2}$ particles gedanken experiment of Greenberger, Horne, and Zeilinger (1989) discussed in chapter 7. This argument is clarified in an article by Clifton, Pagonis, and Pitowsky (1992). See also Herbut (1992) and Pitowsky (1992) for further discussion on Pitowsky's paper.

We will discuss a way of escaping the contradiction suggested by Clifton and Niemann (1992) which involves violating the condition for the existence of elements of reality under certain circumstances. We find that, from a formal point of view, their escape is successful. However, we will find that the class of realistic theories that allow this escape have rather anomalous properties. In particular, they violate a very compelling notion of the continuity of elements of reality. We will consider other ways of escaping the contradiction either by violating the condition for the existence of elements of reality or by violating the condition for the Lorentz invariance of elements of reality. Finally we will show how the contradiction can be illustrated if we assume that the particles actually have trajectories in certain situations.

9.2 The elements of reality

In chapter 4 we reviewed the argument of Einstein, Podolsky, and Rosen for the incompleteness of quantum mechanics. EPR defined a condition for the existence of elements of reality. In the following argument we shall use a similar condition. However, here it is necessary that we are more explicit about what is meant by 'elements of reality'. If we regard reality as being made up of elements then what exactly are these elements? We can be certain that those things we actually measure are elements of reality. For example if we measure a physical quantity Q and it is found to have the value q then we can of course be certain that the physical quantity Q has the value q at the time of measurement. That is we can regard $Q = q$ as an element of reality. In the absence of a measurement we cannot be so certain. Indeed, we cannot then even be sure that it makes any sense to talk of Q as actually having a value. There is, however, one situation in which it would generally be agreed that Q actually has a value even though we have not measured it. This is when we can predict with certainty what the result of making a measurement of Q would be. Thus we have the following *sufficient condition for the existence of an element of reality*:

The physical quantity Q actually has the value q at a given time t if we can predict with certainty (i.e. with probability equal to 1) that the result of measuring Q at time t would be q.

If we can establish that a physical quantity actually has a value then we have established the existence of an element of reality. The reality condition is regarded as 'sufficient' because we may still want to regard Q as actually having a value even when this cannot be established by using this condition. There is a difference in emphasis between this reality condition and that of EPR. Here, we will be more concerned with the actual value of the physical quantity than EPR. Furthermore, there is no phrase referring to not disturbing the system. The latter has more to do with the locality assumption implicit in EPR's argument. Redhead (1987) has defined a reality condition that is equivalent to that above but more closely resembles the EPR reality condition: His condition is:

If we can predict with certainty, or at any rate with probability equal to one, the result of measuring a physical quantity at time t, then at time t there exists an element of reality corresponding to the physical quantity and having a value equal to the predicted measurement result.

To distinguish those elements of reality whose existence we have inferred simply from the fact that we have measured them from those elements of reality whose existence has been inferred by using the reality condition above we denote the latter by putting square brackets around the relevant symbol. Thus, if we have measured Q and obtained the value q then we will put $Q = q$ but if we can predict with certainty that a measurement of Q would give the value q but we have not actually made the measurement then we will put $[Q] = q$.

The reality condition can be illustrated by a macroscopic example. Take a ball and two cups (cup A and cup B). The ball is placed into one of the cups and then the other cup is placed upside down on top of the first cup. Keeping the two cups held together, they are shaken and then they are quickly placed upside down on the table without looking to see which cup the ball has gone underneath. Now, we can lift up one cup, cup A say, to see if the ball is underneath it. If the ball is not underneath cup A then we can predict with certainty that if we were to lift up the other cup then the ball would certainly be underneath it. Now, do we believe that the ball is underneath cup B even if we do not pick the cup up to look? If we believe the reality condition above then we do believe this. There is, of course, no way of proving or disproving the validity of this belief.

In the context of quantum mechanics the reality condition can be stated in the following way:

If, at a given time t, a system is in an eigenstate $|q\rangle$ of an operator $\widehat{Q}$ corresponding to the physical quantity Q, i.e. $\widehat{Q}|q\rangle = q|q\rangle$, then we have $[Q] = q$ at time t even if we do not make a measurement of Q at this time.

Assumption (ii) discussed in chapter 3 is an application of this applied to an atom in a box. We can illustrate this condition with a quantum analogue of the cup and balls experiment discussed above. Consider a single photon impinging onto a beam splitter as shown on a space-time diagram in fig. 9.1. The initial state of the particle is $|s\rangle$. After passing through the beam splitter the state becomes

$$|\psi\rangle = \frac{1}{\sqrt{2}}\Big(i|a\rangle + |b\rangle\Big) \tag{9.1}$$

where $|a\rangle$ ($|b\rangle$) is the state of the particle when it is in path a (b). A detector is placed

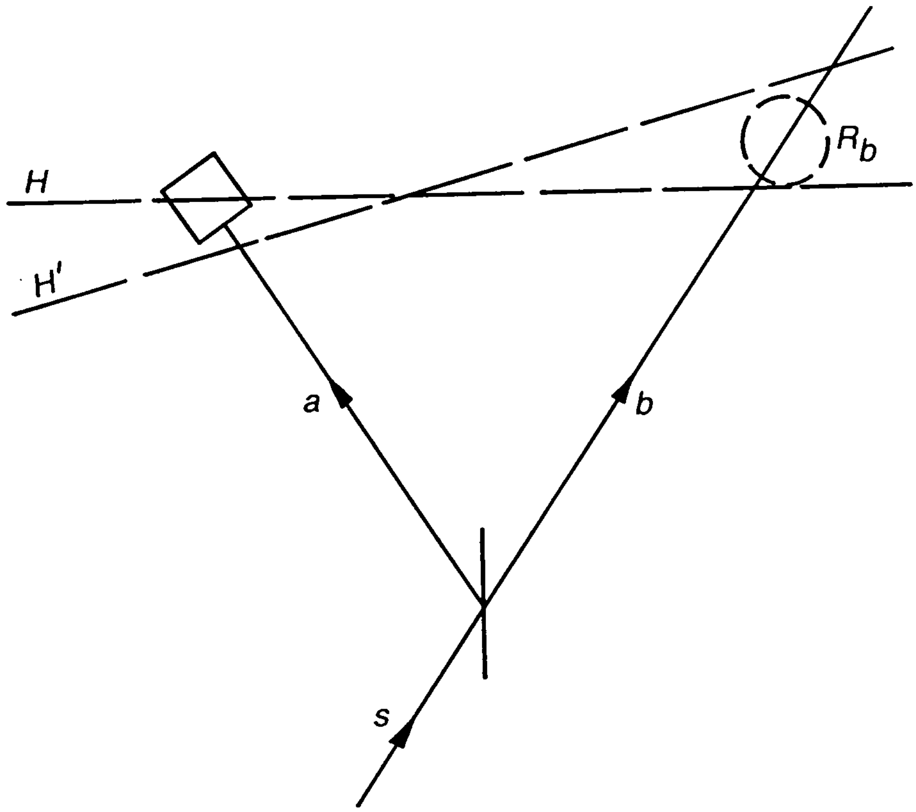

Fig. 9.1. Space-time diagram of a single particle impinging on a beam splitter.

in path a. If the particle is not detected at this detector then the state of the system is projected on to the second term in equation (9.1). That is, the state becomes $|b\rangle$. We consider the operator

$$\widehat{B} = |b\rangle\langle b| \tag{9.2}$$

This corresponds to the physical quantity B which is equal to 1 if the particle is in path b and equal to 0 if the particle is not in path b. The state of the system, $|b\rangle$, is an eigenstate of $\widehat{B}$ with eigenvalue 1 and consequently a measurement of B will certainly yield the value 1. Therefore, by the reality condition, we have that $[B] = 1$ even if such a measurement is not performed. That is, if we believe the reality condition then we believe that the particle is actually in path b when it is not detected in path a even when no measurement is performed to see if it is in path b.

9.3 Lorentz invariance of elements of reality

We assume that we know when the single particle in fig. 9.1 is emitted from the source. Therefore we know what time to expect its detection at the detector in path a if it takes this path. We will call this time t_1. If the particle is not detected in path a at time t_1 then, by the reality condition we know that the particle passes through the space-time region R_b (intersected by path b) shown in fig. 9.1. Here we are looking at the apparatus from the rest frame. In this frame of reference $t = t_1$ is represented by the hyperplane H shown in fig. 9.1. If instead we look at the apparatus from a frame of reference moving to the right then, if the particle had taken path a, it would have reached the detector after the particle would have passed through the region R_b if it had taken path b as viewed from this frame. (This is made clear by looking at the t = *constant* hyperplane, H', passing through R_b shown in fig. 9.1.) In this frame of reference we cannot use the reality condition to establish whether or not the particle passes through the region R_b because we do not have a measurement result from the detector at the time the particle would be in this region if it had taken path b. However, if the particle does actually pass through the region R_b then this cannot depend on the frame of reference we choose to look at the apparatus from. The choice of frame of reference is arbitrary and cannot influence what actually happens. Therefore if we can establish that the particle does actually pass through the region R_b in one frame of reference then this must be true in all frames of reference.

The physical quantity we have considered, whether or not a particle passes through a particular space-time region, is a Lorentz invariant. That is it does not depend on frame of reference of the observer. For our purposes it will be sufficient only to consider such Lorentz-invariant quantities. The above discussion about the particle in one of two paths motivates the following *necessary condition for the Lorentz invariance of elements of reality*:

If we can infer that a Lorentz-invariant physical quantity Q has a value q by applying the reality condition to a physical theory in one frame of reference then the physical quantity has the value q in all frames of reference.

This condition can be generalized to include non-Lorentz-invariant physical quantities by stating that if we can infer that Q has the value q in one frame of reference, F, then it has the value q' in any other frame, F', where q' is the Lorentz-transformed value of q in frame F'. However, in the following we will only be discussing Lorentz-invariant physical quantities. We will regard interpretations in which the reality condition and the

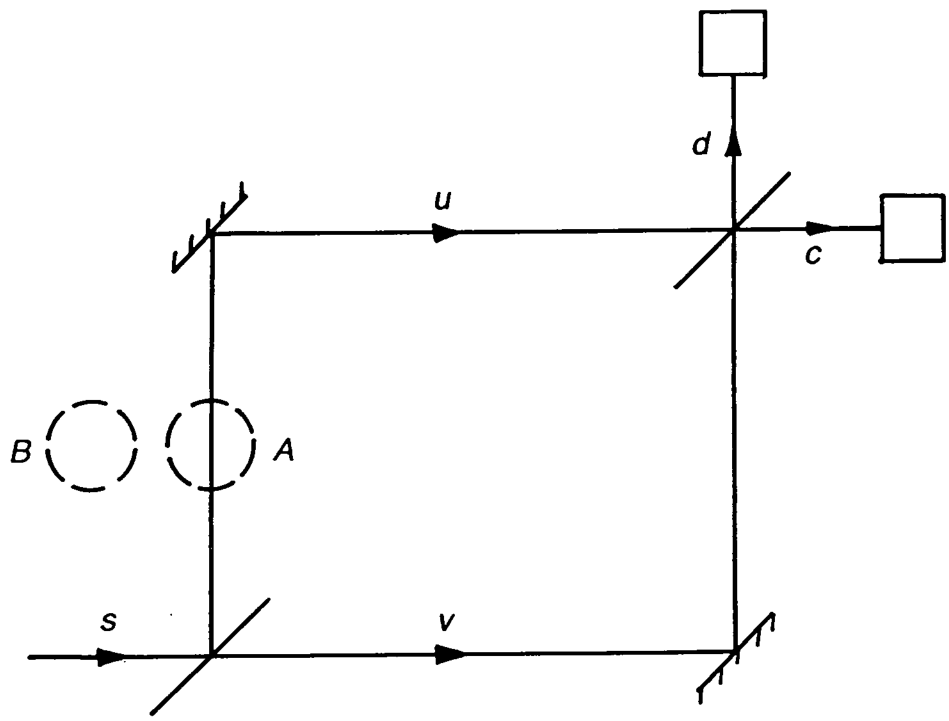

Fig. 9.2. The apparatus discussed by Elitzur and Vaidman. The path difference is set such that when there are no obstacles in the paths there is destructive interference in the d output.

condition for the Lorentz invariance of elements of reality are true as *Lorentz-invariant realistic interpretations.*

Although the observer moving to the right cannot use the reality condition to establish that the particle passes through R_b as can the observer who is at rest, he does not violate the reality condition when he invokes the condition for the Lorentz invariance of elements of reality to infer that the particle passes through the region R_b because the reality condition is only regarded as a sufficient condition. Furthermore, for the example of the particle illustrated in fig. 9.1 the conjunction of the reality condition and the condition for the Lorentz invariance of elements of reality does not lead to any contradictions. It would be very surprising if it did lead to contradictions as there is nothing particularly 'quantum' about this example. It is a direct analogue of the ball and two cups experiment. However, we will see that these conditions do lead to a very striking contradiction when applied to another experimental arrangement.

9.4 The overlapping interferometers gedanken experiment

In the previous chapter we saw how the two overlapping interferometers apparatus was equivelent to the experiment considered in chapter 3 to discuss empty waves. That experiment was motivated by the geganken experiment of Elitzur and Vaidman. For the purposes of this chapter it will be beneficial to consider in more detail the relationship between the Elitzur and Vaidman experiment and the two overlapping interferometers experiment. First, we will recall the details of Elitzur and Vaidman's experiment. We take a Mach-Zehnder-type interferometer for single particles. This interferometer has an input s, two internal paths u and v, and two outputs c and d (see fig. 9.2). The lengths of the two internal paths are set to be equal such that if there are no obstacles in the internal paths then no particles will be detected in output d due to destructive interference. Now, if one of the internal paths is blocked, the u path say, then there will no longer be destructive interference and it will now be possible for particles to emerge at output d. However, if the u path is blocked then any particle being detected in the d output must have reached the detector by going along the v path. Let us imagine that we do not know whether or not the obstacle has been placed into the u path. We can now infer the presence of the obstacle in path u if the particle is detected in the d output even though the particle did not touch the obstacle. The obstacle could be a quantum object. Elitzur and Vaidman consider the case of an atom which is in one of two regions A and B. The atom is prepared in the state

$$\frac{1}{\sqrt{2}}\Big(|A\rangle + |B\rangle\Big) \tag{9.3}$$

where the state $|A\rangle$ ($|B\rangle$) corresponds to the atom being in region A (B). Region A is in path u. If the atom is in region A then we assume that it will block the particle from going along this path. Thus, if the particle is detected in output d then we know that the state of the atom is $|A\rangle$.

As in the previous chapter, we take two Mach-Zehnder type interferometers one for positrons (MZ_1) and one for electrons (MZ_2). Each interferometer has an input s_i ($i = 1, 2$), two internal paths u_i and v_i and two outputs c_i and d_i. Each interferometer also has two 50–50 beam splitters BSA_i and BSB_i. The internal path difference of each interferometer is set equal to 0 such that when each interferometer is taken separately there will be no particles detected in the d_i outputs due to destructive interference. A positron and an electron are created simultaneously and fed into their respective interferometers. The two interferometers are arranged with the u_1 path and the u_2 path overlapping such that if the positron takes path u_1 and the electron takes path u_2 then they will meet at the intersection (point P) and annihilate with probability equal to 1 (see fig. 9.3).

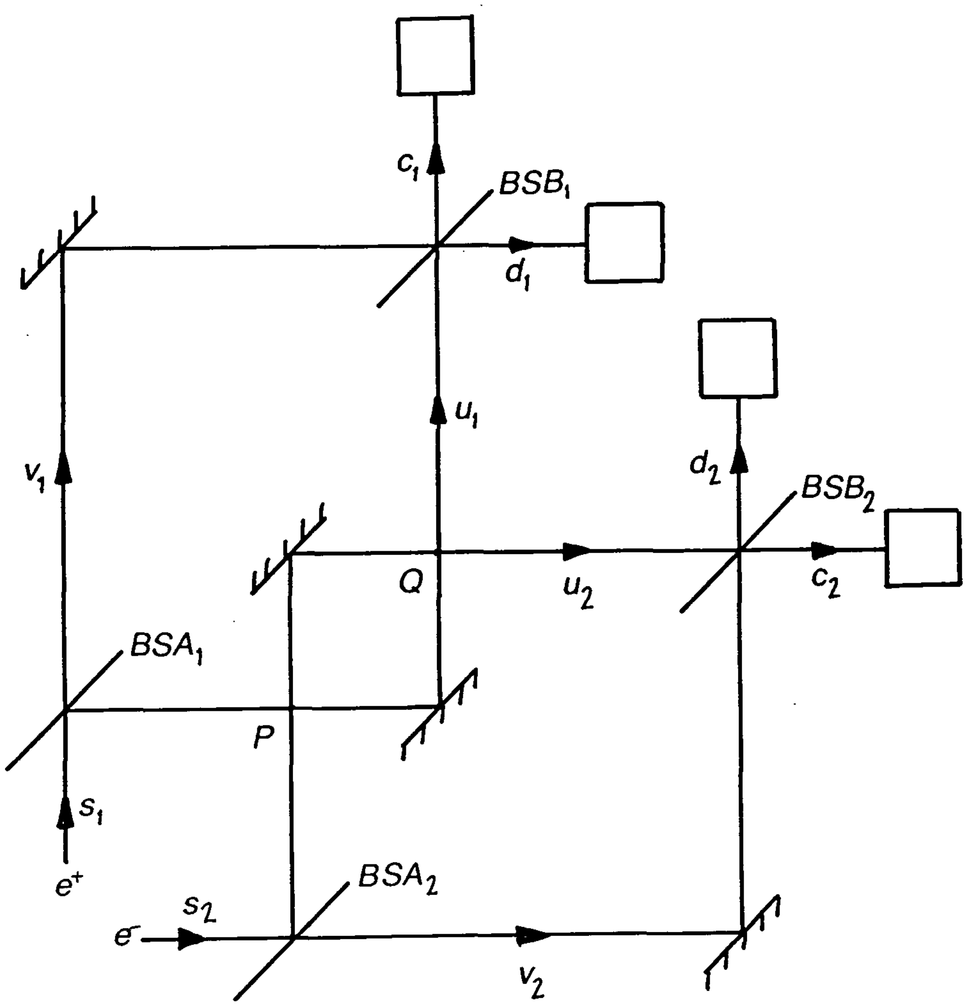

Fig. 9.3. Two overlapping interferometers, one for positrons and one for electrons.

We will consider the physical quantities D_i and U_i with corresponding operators

$$\widehat{D}_i = |d_i\rangle\langle d_i| \tag{9.4}$$

$$\widehat{U}_i = |u_i\rangle\langle u_i| \tag{9.5}$$

These operators have eigenvalues 1 corresponding to the particle being in the relevant path and 0 corresponding to the particle not being in this path.

Now, let us consider the apparatus in a frame of reference F_1 in which the positron passes through BSB_1 (assuming it is not annihilated) and is detected in one of the outputs of MZ_1 before the electron reaches BSB_2. This is now similar to the gedanken experiment of Elitzur and Vaidman with the electron is playing the role of the atom: If the electron takes path u_2 then this effectively blocks path u_1 for the positron. If the positron is detected in the d_1 output (i.e. $D_1 = 1$) then the electron must be in the state $|u_2\rangle$ because otherwise there would be destructive interference in the d_1 output and the positron could not then have been detected in the d_1 output. We have not made a direct measurement on the electron to see if it is actually in path u_2 but by using the reality condition we can conclude that it is actually in this path because its state is $|u_2\rangle$. Thus by applying quantum theory in the frame of reference F_1 we have established that

$$[U_2] = 1 \qquad \text{if} \qquad D_1 = 1 \tag{9.6}$$

It is important to emphasize that we have not actually made a measurement of U_2 but quantum mechanics predicts that if we did then we would certainly find that $U_2 = 1$. Given that the state of the electron is $|u_2\rangle$ we see from (8.3) that there is now a 50% chance that the electron will be detected in the d_2 output after passing through BSB_2. Thus we conclude that

$$D_1 = 1 \quad \text{and} \quad D_2 = 1 \quad \text{sometimes} \tag{9.7}$$

We can derive the results (9.6) and (9.7) by considering the evolution of the quantum state of the system in frame F_1. The initial state of the system is

$$|s_1\rangle|s_2\rangle \tag{9.8}$$

After both the positron has passed through BSA_1 and the electron has passed through BSA_2 the state of the system becomes, using (8.2),

$$\frac{1}{2}\Big(i|u_1\rangle + |v_1\rangle\Big)\Big(i|u_2\rangle + |v_2\rangle\Big). \tag{9.9}$$

After passing point P the state becomes, using (8.1),

$$\frac{1}{2}\Big(- |\gamma\rangle + i|u_1\rangle|v_2\rangle + i|v_1\rangle|u_2\rangle + |v_1\rangle|v_2\rangle\Big). \tag{9.10}$$

After the positron has passed through BSB_1 but before the electron has passed through BSB_2, the state becomes, using (8.3) and (8.4),

$$\frac{1}{2\sqrt{2}}\Big(- \sqrt{2}|\gamma\rangle - |c_1\rangle|u_2\rangle + 2i|c_1\rangle|v_2\rangle + i|d_1\rangle|u_2\rangle\Big). \tag{9.11}$$

If the positron is detected in the d_1 output then the state is projected onto the last term in (9.11) such that the state becomes

$$|d_1\rangle|u_2\rangle \tag{9.12}$$

Therefore the state of the electron is $|u_2\rangle$ as we deduced previously and (9.6) follows from this. After the electron has passed through BSB_2 the state of the system becomes, using

(8.3),

$$\frac{1}{\sqrt{2}}|d_1\rangle\Big(|c_2\rangle + i|d_2\rangle\Big) \tag{9.13}$$

The probability that the positron is detected in output d_1 is equal to $\frac{1}{8}$ (from (9.11)) and the probability that the electron is detected in the d_2 output given that the positron has been detected in the d_1 output is equal to $\frac{1}{2}$ from (9.13). Therefore, the probability that $D_1 = 1$ and $D_2 = 1$ is equal to $\frac{1}{16}$. The important point is that this probability is non-zero. This is result (9.7) above.

We will now consider a frame of reference F_2 in which the electron passes through BSB_2 (assuming it is not annihilated) and is detected in one of the outputs of MZ_2 before the positron reaches BSB_1. The apparatus is symmetrical and therefore, we can use a similar analysis to that which we used in the frame F_1. If the electron is detected in the d_2 output then the state of the positron must become $|u_1\rangle$ such that the u_2 path is blocked to the electron otherwise there would be destructive interference in the d_2 output and the electron could not then have been detected in the d_2 output. Again, this is similar to the gedanken experiment of Elitzur and Vaidman but this time the positron is playing the role of the atom. We have not made a measurement to confirm that the positron is actually in path u_1 but by invoking the reality condition we can conclude that it is. That is, we have

$$[U_1] = 1 \qquad \text{if} \qquad D_2 = 1 \tag{9.14}$$

As the state of the positron is $|u_1\rangle$ we can see from (8.3) that there is a now 50% chance that it will we detected in the output d_1 after passing through BSB_1. Thus we recover the result (9.7) above.

Finally, we consider the rest frame in which the positron and electron impinge on the beam splitters BSB_1 and BSB_2 at the same time. In this frame the state of the system after it has passed point P is the same as that in (9.10). We notice that this state does not contain a $|u_1\rangle|u_2\rangle$ term. Therefore it is an eigenstate of the operator

$$\widehat{U}_1\widehat{U}_2 = |u_1\rangle|u_2\rangle\langle u_2|\langle u_1| \tag{9.15}$$

with eigenvalue 0. Consequently, we have

$$[U_1U_2] = 0 \tag{9.16}$$

in the rest frame. This means that it is not possible for both the positron to be in path u_1 and the electron to be in path u_2 after point P. This is not surprising as if the positron and electron had taken these paths then they would have met and annihilated at point P.

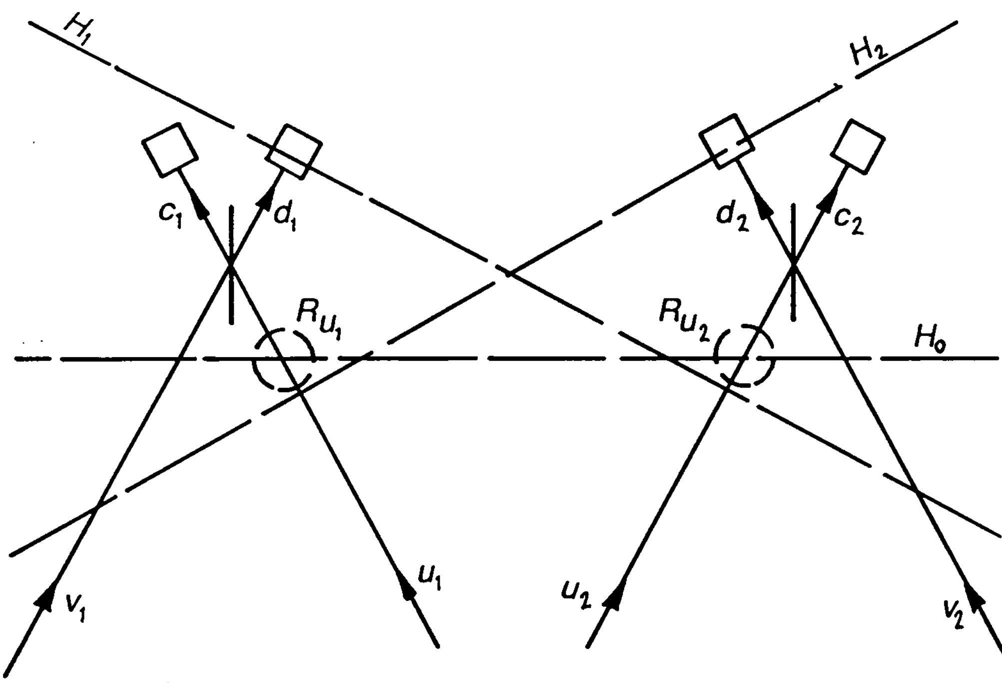

Fig. 9.4. Space-time diagram of part of the apparatus with constant time hyperplanes marked for the F_1, F_2 and rest frames.

9.5 The contradiction

We will now show that if we now apply the condition for the Lorentz invariance of elements of reality we obtain a contradiction. Part of the apparatus is represented on a space-time diagram in fig. 9.4. Let H_1 be a hyperplane that passes through the detection event of the positron in the d_1 output and through path u_2 as shown in fig. 9.4. This hyperplane corresponds to t =constant in the frame F_1. We consider a small region, R_{u_2}, of space-time intersected by the u_2 path after the hyperplane H_1 (see fig. 9.4). We define the physical quantity $U_2(R_{u_2})$ to be equal to 1 if the electron passes through the region R_{u_2} and equal to 0 if the electron does not pass through this region. By invoking the reality condition we see from (9.6) that we can establish that the electron actually passes through R_{u_2} if the positron is detected in output d_1 because this region is after the detection event as viewed from frame F_1. Furthermore, by invoking the condition for the Lorentz invariance of elements of reality, we conclude that the electron passes through this region in all frames of reference. That is we have

$$[U_2(R_{u_2})] = 1 \qquad \text{if} \qquad D_1 = 1 \tag{9.17}$$

in all frames of reference. Now we consider a hyperplane H_2 that passes through the detection event of the electron in the d_2 output and passes through path u_1 as shown in fig. 9.4. This hyperplane corresponds to t =constant in the F_2 frame. Let R_{u_1} be a small region of space-time intersected by path u_1 after the hyperplane H_2. We define the physical quantity $U_1(R_{u_1})$ to be equal to 1 if the positron passes through the region R_{u_1} and equal to 0 if the positron does not pass through this region. By invoking the reality condition we see from (9.14) that the positron actually passes through the region R_{u_1} if the electron is detected in the d_2 output. Furthermore, by invoking the condition for the Lorentz invariance of elements of reality we see that the positron actually passes through this region in all frames of reference. That is

$$[U_1(R_{u_1})] = 1 \qquad \text{if} \qquad D_2 = 1 \tag{9.18}$$

in all frames of reference. Now, from (9.7) we see that sometimes the positron is detected in the d_1 output and the electron is detected in the d_2 output. When this happens we can conclude that the electron passes through R_{u_2} (from (9.17)) and that the positron passes through R_{u_1} (from (9.18)). That is, the physical quantity $U_1(R_{u_1})U_2(R_{u_2})$ has the value

$$[U_1(R_{u_1})U_2(R_{u_2})] = 1 \tag{9.19}$$

in all frames of reference when $D_1 = 1$ and $D_2 = 1$. However, we will see that this is contradicted by the observer in the rest frame. Consider the hyperplane H_0 which passes through the regions R_{u_1} and R_{u_2} as shown in fig. 9.4. By invoking the reality condition we see from (9.16) that it is not possible to have both the positron passing through region R_{u_1} and the electron passing through the region R_{u_2} (as they would have annihilated at point P before reaching these regions). By invoking the condition for the Lorentz invariance of elements of reality we see that this is true for all frames. That is, the physical quantity $U_1(R_{u_1})U_2(R_{u_2})$ always has the value

$$[U_1(R_{u_1})U_2(R_{u_2})] = 0 \tag{9.20}$$

in all frames of reference. This contradicts (9.19). Therefore Lorentz-invariant realistic interpretations of quantum theory are not possible.

9.6 The escape of Clifton and Niemann

Clifton and Niemann have demonstrated that the above contradiction can be avoided if we suppose that the reality condition does not apply to nonlocal physical quantities when the actual measurement context is not appropriate to making a measurement of the nonlocal physical quantity concerned. By a nonlocal physical quantity, we mean that is not associated with only one small space-time region. The reason for picking out nonlocal physical quantities in this way has to do with the nonlocality of quantum mechanics. We must allow the possibility that the value of $U_2(R_{u_2})$ depends on what is measured on the positron because of this nonlocality. Therefore, we need to rewrite (9.17) as

$$[U_2(R_{u_2})]_{D_1} = 1 \qquad \text{if} \qquad D_1 = 1 \tag{9.21}$$

where the subscript D_1 denotes that D_1 is being measured on the other particle. Similarly, we rewrite (9.18) as

$$[U_1(R_{u_1})]_{D_2} = 1 \qquad \text{if} \qquad D_2 = 1 \tag{9.22}$$

Therefore, (9.19) becomes

$$[U_1(R_{u_1})U_2(R_{u_2})]_{D_1,D_2} = 1 \tag{9.23}$$

If keep the reality condition as it is stated in section 9.2 then we still have a contradiction between (9.23) and (9.20) because, when applied in the rest frame, the reality condition states that $U_1(R_{u_1})U_2(R_{u_2})$ is equal to 0 even if it is not measured but instead something else is measured; D_1 and D_2 in this case. However if, as Clifton and Niemann suggest, we modify the reality condition such that, for nonlocal physical quantities, it can only be applied when the context is appropriate to measuring the nonlocal physical quantity concerned then we cannot deduce that $[U_1(R_{u_1})U_2(R_{u_2})]_{D_1,D_2} = 0$ in the rest frame as we did above and thereby we can avoid the contradiction. The context of our experiment is inappropriate to measuring $U_1(R_{u_1})U_2(R_{u_2})$ because if we did actually measure $U_1(R_{u_1})$ and $U_2(R_{u_2})$ by placing detectors in the paths in regions R_{u_1} and R_{u_2} then the results (9.17) and (9.18) would no longer hold.

Although this way of avoiding the contradiction is successful from a formal point of view it does require believing that (9.23) is true. However (9.23) states that the positron actually passes through region R_{u_1} and that the electron actually passes through region R_{u_2}. This is difficult to believe because it seems to require that both the positron and the electron passed through point P without annihilating or alternatively that at least one of the particles jumped into the u_i path after point P. Thus, interpretations that allow Clifton and Niemann's escape must have some rather anomalous properties. In particular they must violate the following *condition for the continuity of elements of reality*:

If a physical quantity Q is measured and found to have the value q then Q continues to have the value q for as long as the evolution of the system is such that the physical quantity Q is conserved.

To see that this condition rules out Clifton and Niemann's escape we consider actually measuring the physical quantity $U_1 U_1$ at a time earlier than the space-time regions R_{u_i} but without measuring U_1 and U_2 separately. We can do this by surrounding the second point where the paths u_1 and u_2 overlap (point Q, see fig. 9.3) by a detector to detect any annihilation products but in such a way that the paths are not blocked. We assume that the positron would arrive at point Q at the same time as the electron and consequently if any annihilation products are detected then $U_1 U_2 = 1$ and if no annihilation products are detected then $U_1 U_2 = 0$. Of course, we will never actually get annihilation products detected as the positron and electron would already have annihilated at point P if they had taken these paths. Therefore, we will always obtain the measurement result $U_1 U_2 = 0$. This measurement does not affect the validity of the results (9.21) and (9.22) and consequently we can still derive the result (9.23). The physical quantity $U_1 U_2$ is conserved until the particles reach beam splitters BSB_i because during this time we have $\frac{d}{dt}\langle U_1 U_2\rangle = 0$ and we know that $U_1 U_2 = 0$ at Q and $U_1 U_2$ cannot be negative. Therefore by invoking the continuity condition we obtain

$$[U_1(R_{u_1})U_2(R_{u_2})]_{D_1,D_2} = 0 \tag{9.24}$$

from the fact that we have $U_1 U_2 = 0$ at Q. The subscripts D_1 and D_2 can remain because the measurement of $U_1 U_2$ at Q is not incompatible with the measurements of D_1 and D_2, i.e. the results (9.21) and (9.22) remain valid when $U_1 U_2$ is measured in this way. Clearly (9.24) contradicts (9.23).

9.7 Violating the reality condition

We will discuss three ways of escaping the contradiction by violating the reality condition defined in section 9.2.

The reality condition requires an ontological commitment so far as the nature of the elements of reality is concerned. Thus, if $[U_1] = 1$ for example then this means that the positron actually passes through R_{u_1}. In a hidden-variable interpretation this ontological commitment is not necessary. We simply require that there exist hidden variables that will serve to determine the outcome of a measurement of $U_1(R_{u_1})$ say. These hidden variables need not necessarily correspond to a description of a positron passing through the region R_{u_1}. Therefore, it is conceivable that a hidden-variable interpretation may violate the reality condition and thereby avoid the contradiction. However, it is shown by Hardy and Squires (1992) that this is not possible for deterministic hidden-variable interpretations in which the hidden-variable descriptions of two subsystems are disjoint when the state of the combined system can be written as a product.

An alternative way of avoiding the contradiction is to abandon the idea of having any elements of reality except for the results of measurements. A physical quantity would only be regarded as having a value when it has actually been measured. It is clear that this approach avoids the contradiction because the reality condition played a crucial role in the above arguments. For measurements that have actually been performed quantum theory is Lorentz invariant. For example, it is not possible to discover a preferred frame of reference by performing quantum mechanical experiments. However, if we take this route then we must regard reality as consisting only of the results of measurements. Quantum mechanics then only serves as a theory that relates the results we see on various macroscopic measurement apparatuses. It is then difficult to see how quantum mechanics can be regarded as a theory that explains what is happening at the microscopic level.

Rather than abandoning the idea of having a reality condition altogether we could consider modified versions of the reality condition. One example is that of Clifton and Niemann discussed above. Another is the following:

A physical quantity Q that can be measured in the space-time region R actually has the value q in this region when observers at all points in R can have information that will allow them to predict with certainty that the result of measuring Q in this region will be q.

An observer can only have information about events in his backward light cone. Consequently, if the positron is detected in the d_1 output then we cannot infer that $[U_2(R_{u_2})] = 1$ by this reality condition because the detection of the positron does not happen in the backward light cone of an observer in the region R_{u_2}. At first sight this reality condition may seem like the most obvious way of avoiding the contradiction. It is clearly formulated in a Lorentz covariant way. However, from the point of view of realism, it is quite problematic. Consider the example of the cup and two balls of section 9.2. If the ball is found not to be under cup A then where is it? The above reality condition does not give an answer to this question until information about the measurement on cup A has reached cup B. As the reality condition is a sufficient condition we could say that the ball is actually under cup B anyway by way of trying to answer the question. However, this takes us back to the original reality condition in section 9.2. These comments also apply to the quantum mechanical example of the particle impinging onto a beam splitter shown in fig. 9.1.

9.8 Violating the Lorentz-invariance condition

The condition for the Lorentz invariance of elements of reality defined in section 9.3 contains two parts

(a) The physical theory can be applied in any frame of reference when it is being used to infer the value of a physical quantity by using the reality condition.

(b) Having inferred that a Lorentz invariant physical quantity has a value in one frame of reference we assume that is has this value in all frames of reference.

We can avoid the contradiction by violating either of these two parts. First consider violating part (b). We could replace (b) by:

(b$'$) If we can infer that a Lorentz-invariant physical quantity Q has the value q with respect to one frame of reference F then it does not necessarily have a value with respect to another frame of reference F' or if it does have a value then this value is not necessarily the same.

It is clear that we can avoid the contradiction if we replace (b) by (b$'$) because we cannot then compare the results we have inferred by using the various different frames of reference. However, (b$'$) is unsatisfactory because allows different things to happen in different frames of reference. This clearly violates Lorentz invariance and it is also difficult to see how this can be understood from a realist point of view. For example, it would allow a particle to go along one path in one frame of reference and along another path in another frame of reference. This particular problem can be avoided a modification of (b) due to Clifton and Niemann (1992):

(b$''$) If a Lorentz-invariant physical quantity Q has a value q with respect to one frame F then if it also has a value with respect to another frame F' this value is also q.

This avoids the contradiction because we can not infer the existence of an element of reality for region R_{u_1} by invoking the reality condition in frame F_1. However, we will see in the next section that if we assume trajectories in certain circumstances then (b$''$) is violated because we will be able to deduce that the electron, for example, takes path u_2 with respect to frame F_1 and path v_2 with respect to frame F_2. Even without assuming trajectories (b$''$) may still be regarded as unsatisfactory because, as was the case with (b$'$), what happens is allowed to depend on the frame of reference. Thus, we may be able to deduce that a particle passes through a particular space-time region in one frame but in another frame it may be is meaningless to ask whether the particle passed through this region or not. This violates Lorentz invariance in the sense that the choice of frame of reference has an effect on what happens and it goes against the realist view point in that different things happen in different frames.

Instead of violating part (b) of the condition for the Lorentz invariance of elements of reality we could violate part (a) and thereby avoid the contradiction. The way to violate part (a) is not to allow the use of the physical theory in all frames to infer the existence of elements of reality. The derivation of the contradiction employed three frames, F_1, F_2, and the rest frame. The rest frame is not needed if one adopts the condition for the continuity of elements of reality of section 9.6. In this case, we only need two frames of reference to run the argument. These could be any two frames because the rest frame of the apparatus could always be chosen such that the velocity of the two frames is equal and opposite relative to this rest frame. The two frames could then serve as F_1 and F_2. Therefore, if we assume the continuity condition then the only way to avoid the contradiction by violating (a) is to have one preferred frame of reference in the universe. Part (a) then becomes

(a$'$) There exists a preferred frame of reference in the universe and if the physical theory is applied in any frame of reference other than this preferred frame to infer the value of a physical quantity by using the reality condition then we may obtain the wrong value for this quantity (or obtain a value when the physical quantity does not actually have a value).

Any notion of a preferred frame of reference does, of course, violate Lorentz invariance. However, it is not difficult to understand the idea of a preferred frame of reference from the realist viewpoint. Therefore, escaping the contradiction by this route does not run us into the difficulties with realism we have found with the other escape routes. Having inferred the value of a Lorentz-invariant physical quantity by applying the reality condition in the preferred frame, we can then apply part (b) to infer that this physical quantity has the same value in all frames of reference.

9.9 Illustrating the contradiction with trajectories

The violation of Lorentz invariance can be illustrated in a most dramatic way if adopt the following *sufficient condition for the existence of trajectories*:

If a particle is in a particular region R and there is only one path open by which the particle may have reached this region then the particle actually went along this path.

This goes beyond the provisions of the reality condition of section 9.2 and therefore can be regarded as an addition to that reality condition. We will now apply this in each of the F_1 and F_2 frames of reference.

Consider first the F_1 frame. In this frame we can deduce from (9.6) that if the positron is detected in the d_1 output then the electron passes through region R_{u_2}. Therefore, by the trajectory condition the electron must actually go along path s_2 and then along path u_2 to region R_{u_2}. If the electron is subsequently detected in the d_2 output then, again by the trajectory condition it must actually go from region R_{u_2} along the remainder of path u_2 and then along path d_2 until it arrives at the detector. That is, it has the trajectory shown in fig. 9.5(a). Now, if the positron were to go along the u_1 path it would be annihilated as it would meet the electron at point P. Consequently path u_1 is effectively blocked to the positron. Hence, the only path that is open for the positron to have reached the detector in the d_1 output is to go along path s_1 then along the v_1 path and then along the d_1 path. Therefore, by the trajectory condition it must actually have gone along this path. That is, it has the trajectory shown in fig. 9.5(a).

Now consider the F_2 frame. If the electron is detected in the d_2 output then from (9.14) we see that the positron passes through the region R_{u_2}. Therefore, by a similar argument to that used in frame F_1 we can deduce that the positron goes along s_1 then along u_1 and then along d_1 if it is detected in the d_1 output. Also, by the same reasoning as before, we can deduce that the electron goes along s_2 then along v_2 and then along d_2. Thus in frame F_2 we calculate the trajectories of the particles to be those shown in fig. 9.5(b).

Thus, we find that for the same outcome, i.e. when $D_1 = 1$ and $D_2 = 1$, an observer applying the reality condition and the continuity condition in frame F_1 will calculate different trajectories to those calculated by another observer doing the same thing in the F_2 frame. Clearly, the trajectories calculated in the two frames are contradictory. This contradiction is best avoided if we suppose that there is a preferred frame of reference in

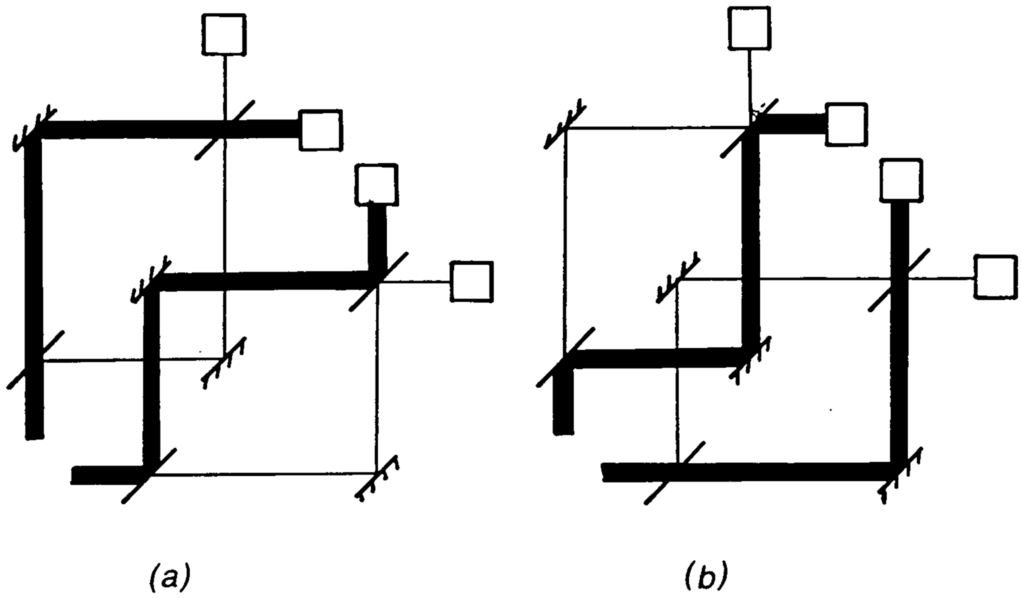

Fig. 9.5. Trajectories calculated (a) in the F_1 frame of reference and (b) in the F_2 frame of reference.

the universe as discussed in the previous section. If this preferred frame is moving in the same direction relative to the rest frame of the apparatus as the F_1 frame say, then the true trajectories actually taken by the particles will be those shown in fig. 9.5(a).

9.10 Conclusions

We have demonstrated that when the condition for the existence of elements of reality and the condition for the Lorentz invariance of elements of reality are applied to quantum mechanics we obtain a contradiction. Therefore, quantum mechanics cannot be given a Lorentz-invariant realistic interpretation. There are various ways of avoiding the contradiction by violating one or the other of the conditions. From a realist view point the most satisfactory way of avoiding the contradiction is to have a preferred frame of reference in the universe. The contradiction is illustrated in a most dramatic way if we assume that particles actually have trajectories in those situations in which there is only one path available to the particle. When we do this we find that contradictory trajectories are calculated in different frames of reference.

Chapter 10

Conclusions

We have examined three main topics in this work: Wave-particle duality, locality, and Lorentz invariance. Each has been considered from a realist point of view although the notion of realism used in each case depended on the context.

In chapter 2 the problem of wave-particle duality was investigated. Here we adopted perhaps the most simplistic notions of realism used in this work. We attempted to explain single particle quantum phenomena (with the example of photons being used) firstly in terms of a classical particle picture alone and then secondly in terms of a classical wave picture alone. The attempt to explain all single particle quantum effects in terms of a particle picture failed because interference effects cannot be explained in this way. It is less straight forward to see that these phenomena cannot be explained in terms of a classical wave picture. The single particle anticorrelation experiment often used to illustrate the particle nature of single particle quantum effects can in fact be explained by classical wave theory if the transmissivity is allowed to vary in a stochastic way. However, such a variation would effect the results of an interference experiment. This is made quantitative by deriving an inequality that allows a direct comparison of an anticorrelation experiment and an interference experiment. If the inequality is violated then classical wave theory cannot explain these phenomena. Quantum mechanics predicts that the inequality will be violated and therefore an appeal to experiment is required to resolve the issue. Whilst a single photon anticorrelation experiment and a single photon interference experiment have been performed by Grangier et al. (1986) the interference experiment they performed was not of the type necessary to test the inequality and consequently it remains an open question as to whether we can explain single particle quantum phenomena in terms of a classical wave picture. However, if we allow some additional assumptions then quantities measured in the experiments performed by Grangier et al. can be substituted into the inequality and it is then found that the inequalities are violated. Therefore, it seems unlikely that classical wave theory can explain single particle quantum phenomena but there does remain some doubt about this conclusion because of the need for the additional assumptions.

If single particle quantum phenomena cannot be explained in terms of particles alone or in terms of waves alone then there are two options open to us. Firstly, we could let the nature of the quantum system depend on the experimental context. Thus, anticorrelation experiments could be explained in terms of particles and interference experiments could be explained in terms of waves. This approach is rather unsatisfactory because the nature of a quantum system is allowed to depend on which experiment is being performed even though we can decide which experiment to perform after the quantum system has already passed through the beam splitter and therefore the quantum system would need to 'know' the future. This is Wheeler's delayed choice experiment. Secondly, we could suppose that both a wave and a particle exist. This avoids the problems of Wheeler's delayed choice experiment and is the approach taken in the de Broglie-Bohm interpretation. Each particle is taken to be accompanied by a wave. When a particle plus wave impinge onto a beam splitter the particle can only go along one of the paths but the wave is divided and part of it goes along each path. The wave that goes along the other path to the particle is empty, that is it does not have a particle in it. In chapter 3 we addressed the question of whether these empty waves can be shown to exist. In one sense, standard interference experiments give evidence for the existence of these empty waves because if the particle only goes along one path then something must go along the other path to extract the path length information such that the appropriate interference pattern will emerge. However, in such experiments we do not know which path the particle goes along and consequently we cannot necessarily assert that there is a particle that only goes along one path. Furthermore, even if we do regard this assertion to be true, the empty wave does not actually manifest its reality in the region where it is empty but rather when it is recombined with the particle. Motivated by these observations it is suggested that sufficient conditions for the existence of empty waves are that we know which path the particle takes and that some system placed in the other path has some measurable property changed. We found that these conditions are satisfied when applied to a particular gedanken experiment if we allow three realist motivated assumptions. The gedanken experiment consists of a box with a spin $\frac{1}{2}$ atom in it which is divided into two boxes and one of these boxes is placed into one of the internal paths of an interferometer. It is found that sometimes we can deduce that the interferometer particle went along the other path to the path with the box in it. Thus, only the empty wave impinges on the box. However, when the boxes are brought back together and a measurement is made on the atom it is found that sometimes there has been a change in the value of the measurable property. Therefore, the conditions for the existence of empty waves are satisfied.

Nonlocality is already apparent in the gedanken experiment for demonstrating the existence of empty waves. In chapters 4 to 8 we considered nonlocality in more detail. In chapter 4 the thought experiment of Einstein, Podolsky, and Rosen to demonstrate the incompleteness of quantum mechanics was discussed. EPR define a complete theory to be one in which every element of physical reality is represented in the physical theory. They also define the following sufficient criterion for the existence of an element of reality: if, without disturbing the system, we can predict with probability equal to one the value of a physical quantity then there exists an element of physical reality corresponding to this physical quantity. They then show how it is possible, by considering a two particle entangled state, to demonstrate the existence of elements of reality corresponding to two physical quantities that correspond to two noncommuting operators. Now, values for physical properties corresponding to two noncommuting operators cannot enter into the quantum mechanical state description of a system and consequently, it is claimed that quantum mechanics is incomplete. However, their argument makes an implicit assumption of locality. This assumption of locality was questioned by Bell. He found that by invoking this assumption of locality and making certain realist assumptions, he was able to derive a set of inequalities which are violated by the predictions of quantum mechanics for the two particle singlet state. This demonstrates that quantum mechanics is not a local realistic theory. In deriving the Bell inequalities the origin of the contradiction between quantum mechanics and local realism is obscured. In chapter 4 we consider a new way of deriving Bell inequalities that does not obscure this contradiction. A set of statements are considered which are clearly contradictory when all true if local realism is assumed. It is found that quantum mechanics predicts that the probability that all these statements are true is greater than zero in certain circumstances, that is there is a nonzero probability of a contradiction.

Bell considered the two particle singlet state when constructing his argument. Most experiments that have actually been conducted to test local realism have involved photons. One problem that these experiments have faced is that the two photons, when emitted form the source, have poor directional correlation. This together with the low efficiencies of the available detectors necessitates the use of supplementary assumptions if a test of local realism is to be possible. Improvements in detector efficiencies requires new technological innovations. However, there seems to be no reason to doubt that sufficiently high efficiency detectors will one day be available. This leaves the problem of directional correlation. In fact one experiment in which the directions of the photons were well correlated has already been performed. This is the double interferometer set up proposed by Horne, Shimony, and Zeilinger (1987) and performed by Rarity and Tapster (1990). In chapter 5 another experiment in which the photons are well correlated in direction is proposed. The two photon state produced is a singlet-type state in that it involves correlations between the polarization of the two photons (like the correlation between the spin of the two particles in the singlet state). The fact that such states can actually be produced demonstrates that the conjecture of Santos (1991) that it is not possible to realise the conditions for a violation of the Bell inequalities without supplementary assumptions even if ideal detectors were available is false. However, the issue of supplementary assumptions will not be finally resolved until an experiment is performed which renders them redundant.

Usually, in experiments to test Bell's inequalities we consider a source of particle pairs. However, Tan, Walls, and Collett demonstrated that Bell's inequalities could be violated by a source of single photons. The single photon impinges onto a beam splitter and homodyne detection is performed on each output from this beam splitter. Unfortunately, as we saw in chapter 6, this proposal requires a supplementary assumption even in the case of an ideal experiment. We considered another proposal to demonstrate the nonlocality of a single photon. In this experiment an atom plus boxes system was placed in each path of a Mach-Zehnder interferometer for single photons. It was found that if the photon was detected in one output of the interferometer then the states of the two atoms become entangled even though they only interacted via a single photon. Whilst this may be regarded as demonstrating the nonlocality of a single photon there is a problem. This is that the two paths open to the photon come back together and therefore the output the photon goes into at the second beam splitter of the interferometer can be influenced by the hidden-variable descriptions of the atoms. That is, the selection of the subensemble considered depends on three particles (two atoms and one photon). Therefore, whilst local realism is certainly violated (without the need for supplementary assumptions), there remains some question as to whether this effect can truly be regarded as demonstrating the nonlocality of a single photon. However, in a third proposal, we find that by combining the first two proposals, we are able to overcome the problems with each of them. In this problem-free proposal the single photon impinges onto a beam splitter, then an atom plus boxes system followed by a homodyne detector is placed in each of the two outputs from this beam splitter. By considering only certain homodyne detection events we find that the state of the two atoms becomes entangled and we are able to get a violation of the Bell inequalities. Therefore, we can truly demonstrate the nonlocality of a single photon.

We go from considering one photon in chapter 6 to considering three or more particles in chapter 7. Greenberger, Horne, and Zeilinger's demonstration of Bell's theorem using three or more particles was reviewed. Their demonstration has two new features. First, it does not require inequalities and secondly it applies to every run of the experiment. However, in a real experiment inequalities would be required as we would not expect to obtain the perfect correlations of gedanken experiments. In chapter 7 it is shown how the methods of chapter 4 can be extended to derive N-particle inequalities. We find that when there is a maximum possible violation of these inequalities an inequality-free demonstration of Bell's theorem is possible. This establishes the connection between the inequalities approach and the GHZ approach to Bell's theorem. Another way of seeing this connection is by reference to the new way to obtain Bell inequalities discussed in sections 4.4 and 7.4. We can write down a number of statements that apply to the N-particle state which, when all true, are contradictory if local realism is assumed. Inequalities are necessary if the probability that they are all true is less than one. However, inequalities are not necessary if the probability that they are all true is equal to one. We find that this probability can only be equal to one for three or more particle states. The exception to this is when the number of settings of the local variables tend to infinity as discussed in section 4.4.

Another way to demonstrate Bell's theorem without inequalities, this time for only two particles and only two settings of the local variable for each particle was presented in chapter 8. The gedanken experiment considered, which is equivalent to the gedanken experiment considered in chapter 1 to discuss empty waves, consists of two overlapping interferometers. The strength of this proposal lies mainly in its simplicity. The state considered is dissimilar to the singlet state considered by Bell and the N-particle states considered by GHZ in that there is not perfect correlation between appropriate measurements on the particles. However, there is a result dependent perfect correlation and consequently, it is possible to apply EPR type reasoning to infer the existence of the result functions necessary to derive the contradiction between local realism and quantum mechanics. Another gedanken experiment was proposed which is the analogue of the two overlapping interferometers set up but which involves two atom plus boxes systems. However, neither of these experiments is sufficiently practical that they could be performed in a laboratory. Thus, a realizable quantum optical version of the two overlapping interferometers was also proposed in section 8.6. In a real experiment we could not expect to obtain the result dependent perfect correlation necessary to derive the theorem without inequalities. Therefore inequalities are necessary. These inequalities were derived. Thus, all the necessary theoretical work has been done to make an experiment possible.

In chapter 9 we saw how, by defining a sufficient condition for the existence of elements of reality similar to that of EPR and also defining a necessary condition for the Lorentz invariance of these elements of reality, we were able to obtain a contradiction with quantum mechanics. This demonstrates that quantum mechanics is not a Lorentz invariant realistic theory. To demonstrate this we considered the two overlapping interferometers gedanken experiment also used in chapter 8. The contradiction with Lorentz invariance is illustrated in a most dramatic way if we allow the existence of trajectories at least in those situations for which there is only one path available to the particle. When we do this we find that different trajectories are calculated in different frames of reference. From a realist point of view the best way of avoiding the contradiction is to have a preferred frame of reference in the universe. When quantum mechanics and the reality condition are applied in any other frame then they will give the wrong answers in some situations. This frame of reference can be understood as the one in which the nonlocal effects are instantaneous.

It remains a matter of opinion as to whether the realist approach to quantum mechanics is the right one. However, if we take such an approach then we can demonstrate the existence of empty waves and that locality and Lorentz invariance are violated.

Bibliography

Alley C. O. and Shih Y. H. (1987) in Proc. 2nd Int. Symp. on the foundations f quantum theory in the light of new technology, eds. Namiki et al. (Physical Society of Japan, Tokyo).

Aspect A., Grangier P. and Roger G. (1981) **47**, 460.

Aspect A., Grangier P. and Roger G. (1982) **49**, 91.

Aspect A., Dalibard J. and Roger G. (1982) **49**, 1804.

Aspect A. and Grangier P. (1987), Hyperfine Interactions **37**, 3.

Aspect A. (1990) in Sixty-two years of uncertainity, ed. A. I. Miller (Plenum Press, Ney York).

Ballentine L. E. and J. P. Jarret (1987), Am. J. Phys. **55**, 696.

Bell J. (1964), Physics (Long Island City, N. Y.)**1**, 195.

Bell J. (1966), Rev. Mod. Phys. **38**, 447.

Bernstein H. J., Greenberger D. M., Horne M. A. and Zeilinger A. (1992), A Bell theorem without inequalities for two spinless particles, submitted to Phys. Rev. A.

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

Bohm D. (1952), Phys. Rev. **85**, 166.

Bohm D., Hiley B. and Kaloyerou P. N. (1987), Phys. Rep. **144** 321.

Bohr N. (1935), Phys. Rev. **48**, 696.

Braunstein S. L. and Caves C. M. (1989) in Proc. 3rd Int. Symp. on Foundations of quantum mechanics in the light of new technology, eds. S. Kobayashi et al. (Physical Society of Japan).

Braunstein S.L., Mann A. and Revzen M. (1992), Phys. Rev. Lett. **68**, 3259.

Brown H. R. and G. Svetlichny (1990), Found. Phys. **20**, 1379.

Campos R. A., Saleh B. E. A. and Teich M. C. (1989), Phys. Rev. A **40**, 1371.

Chu S. (1992), Laser trapping of neurtal particles, Sci. Am., Febuary issue.

Clauser J. F. and Horne M. A. (1974), Phys. Rev. D bf 10, 526.

Clauser J. F. (1976), Phys. Rev. Lett. **36**, 1223.

Clauser J. F., Horne M. A., Shimony A. and Holt R. A. (1974), Phys. Rev. Lett. **23**, 880.

Clifton R. K. and P. Nieman (1992), Phys. Lett. A **166**, 177.

Clifton R. K., Pagonis C. I. Pitowsky (1992), relativity, Quantum mechanics and EPR, Pilosophy of Science Association 1992, Vol. I. (to be published). Einstein A., Podolsky B. and Rosen N., (1935), Phys. Rev. **47**, 777.

Clifton R. K., Redhead L. G. and Butterfield J. N. (1991), Found. Phys. **21**, 149.

Elitzur A. C., Popescu S. and Rohrlich D. (1992), Phys. Lett. A **162**, 25.

Elitzur A. C. and Vaidman L.(1991), Quantum mechanical interaction-free measurements, Tel Aviv preprint.

Fine. A (1982), J. Math. Phys. **23**, 1306

Franson J. D. (1989), Phys. Rev. Lett. **67**, 290.

Freedman S. J. and Clauser J. F. (1976), Phys. Rev. Lett. **28**, 938.

Fry E. S. and Thompson (1976), Phys. Rev. Lett. **37**, 465.

Garuccio A., Popper K. and Vigier J. P. (1981), Phys. Lett. A 86, 397.

Garuccio A., Rapisarda V. and Vigier J. P. (1982), Phys. Lett. A 90, 17.

Ghose P., Home D. and Agarwal G. S. (1991), Phys. Lett. A **153** 403.

Ghose P. and Roy M. N. S. (1991), Phys. Lett. A **161**, 5.

Gisin N. (1991), Phys. Lett. A **154**, 201.

Gordon P. E. (1989), Phys. Lett. A **138**, 359.

Grangier P. (1986), These de doctorat d' étet, Université Paris XI, Orsay.

Grangier P.,Roger G. and Aspect A. (1986), Europhys. Lett. **1**, 173.

Grangier P., Potasek M. J. and Yurke B. (1988), Phys. Rev. A **38**, 3132.

Greenberger D. M., Horne M. A. and Zeilinger A. (1989) in Bell's theorem, quantum theory and conceptions of nature, ed. M. Kafatos (Kluwer, Dordrecht).

Greenberger D. M., Horne M. A., Shimony A. and Zeilinger A. (1990), Am. J. Phys. **58**, 1131.

Hardy L. (1991a), Europhys. Lett. **15**, 591.

Hardy L. (1991b), in Proc. 2nd Int. Wigner Symp.

Hardy L. (1991c), Phys. Lett. A **160**, 1.

Hardy L. (1991d), Phys. Lett. A **161**, 21.

Hardy L. (1992e), Phys. Lett. A **161**, 326.

Hardy L. and Squires E. J. (1992), Phys. Lett. A. **168**, no 3. (forthcoming).

Herbut F. (1992), Phys. Lett. A **163**, 5.

Heywood P. and Redhead M. L. G. (1983), Found. Phys. **13**, 481.

Holland P. R., (1992) The quantum theory of motion (Cambridge University Press, Cambridge) to be published.

Holt R. A. and Pipkin F. M. (1974), Harvard Universtiy preprint.

Hong C. K. and Mandel L. (1985), Phys. Rev. A **31**, 2409.

Hong C. K., Ou Z. Y. and Mandel L. (1987), Phys. Rev. Lett. **59**, 2044.

Horne M. A., Shimony A. and Zeilinger A. (1989), Phys. Rev. Lett. **62**, 2209.

Klyshko D. N. (1988), Photons and nonlinear optics, translated by Y. Sviridov (Gordon and Breach, New York).

Kochen S. and E. Specker (1967), J. Math. Mech. **17**, 59.

Lepore V. L. and Selleri F. (1990), Found. Phys. Lett. **3**, 203.

Loudon R. (1980), Rep. Prog. Phys. **43**, 913.

Mandel L. (1984), Phys. Lett. A **103**, 416.

Marshall T. and Santos E. (1987), Europhys. Lett. **3**, 293.

Marshall T. and Santos E. (1988), Found. Phys. **18**, 185.

Marshall T. and Santos E. (1989), Phys. Rev. A **39**, 6271.

Mermin N. D. (1990a), Phys. Rev. Lett. **65**, 1838.

Mermin N. D. (1990b), Phys. Rev. Lett. **65**, 3373.

Mermin N. D. (1990c), Phys. Today, June issue, pg. 9.

Misra B. and Sudarshan E. C. G. (1977), J. Math. Phys. **18**, 756.

Morrow B. R. (1973), Phys. Rev. A **8**, 2684.

Oliver B. J. and Stroud Jr. C. R. (1989) Phys. Lett. A. **135**, 407.

Ou Z. Y. (1607), Phys. Rev. A **37**. 1607.

Ou Z. Y. and Mandel L. (1988), Phys. Rev. Lett. **61**, 50.

Pagonis C. (1992), Empty waves: not necessarily efective, forthcoming in Phys. Lett. A.

Pagonis C. and R. Clifton (1992), Phys Lett. A. **168**, 100.

Pagonis. C. Redhead L. G. and Clifton R. K. (1991), Phys. Lett. A **155**, 441.

Pascazio S. and Reignier J. (1987), Phys. Lett. A **126**, 163.

Pitowsky I. (1991), Phys. Lett. A **156**, 137.

Pitowsky I. (1992), Phys. Lett. A **166**, 292.

Popescu S. and Rohlich D. (1992), Phys. Lett. A **166**, 293.

Rarity J. G. and Tapster P. R. (1990), Phys. Rev. Lett. **64**, 2495.

Redhead M. L. G. (1987), Incompleteness, nonlocality and Realism (Clarendon, Oxford).

Roy S. M. and Singh V. (1991), Phys Rev. Lett. **67**, 2761.

Santos E. (1990), Found. Phys. bf 21, 221.

Santos E. (1991), Phys. Rev. Lett. **66**, 1388.

Selleri F. (1969), Lett. Nuovo Cimento **1**, 591.

Selleri F. (1989) in Quantum theory and pictures of reality, ed. W. Schommers (springer, Berlin).

Squires E. J. (1990), Phys. Lett. A **148**, 381.

Stairs A. (1983), Philos. Sci. **50**, 587.

Stapp H. P. (1980), Found. Phys. **10**, 767.

Stapp H. P. (1990), Interpretation of an experiment of the Greenberger-Horne-Zeilinger kind, Lawrence Berkelel Laboratory, preprint LBL-293377.

Suppes P. and Zanotti M. (1991), Found. Phys. Lett. **4**, 101.

Tan S. M., Walls D. F. and Collett (1991), Phys. Rev. Lett. **66**, 252.

Tarozzi G. Lett. (1985), Nuvo Cimento **42**, 438.

Yurke B. and Stoler D. (1992a), Phys. Rev. Lett. **68**, 1251.

Yurke B. and Stoler D. (1992b), Using the Pauli exclusion principle to exhibit local realism violations in overlapping interferometers, submitted to Phys. Rev. A.

Zou X. Y., Wang L. J. and Mandel L. (1991), Phys. Rev. Lett. **67**, 318.

Zeilinger A. (1986), Phys. Lett. A **118**, 1.

Żukowski M. (1989), Phys. Lett. A **134**, 351.

Żukowski M. (1990), Phys. Lett. A **150**, 136.

Żukowski M. (1991), Phys. Lett. A **157**, 203.

Żukowski M. and Pykacz J. (1988), Phys. Lett. A **127**, 1.

Żukowski M. and Zeilinger A. (1991) Phys. Lett. A **155**, 69.