Bell's theorem without inequalities[a)]

Daniel M. Greenberger
Department of Physics, City College of the City University of New York, New York, New York 10031
Michael A. Horne
Department of Physics, Stonehill College, North Easton, Massachusetts 02357
Abner Shimony
Departments of Philosophy and Physics, Boston University, Boston, Massachusetts 02215
Anton Zeilinger
Atominstitut der Österreichischen Universitäten, Schüttelstrasse 115, A-1020 Vienna, Austria

(Received 10 June 1990; accepted for publication 30 July 1990)

It is demonstrated that the premisses of the Einstein–Podolsky–Rosen paper are inconsistent when applied to quantum systems consisting of at least three particles. The demonstration reveals that the EPR program contradicts quantum mechanics even for the cases of perfect correlations. By perfect correlations is meant arrangements by which the result of the measurement on one particle can be predicted with certainty given the outcomes of measurements on the other particles of the system. This incompatibility with quantum mechanics is stronger than the one previously revealed for two-particle systems by Bell's inequality, where no contradiction arises at the level of perfect correlations. Both spin-correlation and multiparticle interferometry examples are given of suitable three- and four-particle arrangements, both at the gedanken and at the real experiment level.

I. INTRODUCTION

Bell's theorem of 1964 is a proof that certain plausible propositions about locality, reality, and theoretical completeness are incompatible with some predictions of two-particle quantum mechanics.[1] These propositions were presented in 1935 by Einstein, Podolsky, and Rosen (EPR),[2] who used them in conjunction with some quantum mechanical predictions as the premisses of an argument concluding that quantum states cannot in all situations be complete descriptions of physical reality. There is nothing in EPR's argument nor in their comments on it that suggests doubt about the correctness of quantum mechanical predictions. Their claim that the quantum state is an incomplete description is offered rather as an interpretation of quantum mechanics: roughly, that individual systems have intrinsic properties, but the quantum state gives only a statistical description of an ensemble of intrinsically differing individual systems. Furthermore, they suggest that the idea of a complete state, richer in content than the quantum state, provides a commonsense explanation of certain perfect correlations predicted by quantum mechanics, which otherwise are baffling. Consequently, contrary to the impressions of many physicists, EPR were not offering a paradox, but rather a program for solving a problem. What Bell's theorem shows is that their program cannot be right: The conjunction of EPR's propositions with the quantum mechanical predictions for a pair of systems (as simple as two-state systems) leads to a contradiction.

The contradiction is revealed by deriving from EPR's program an *inequality* which is violated by certain quantum mechanical *statistical predictions*. Statistical predictions concern imperfect or statistical correlations, in which the outcome of a measurement on one system determines not the outcome of a measurement on the other system but rather the probabilities of various outcomes. Because of the contradiction, experiments can be performed in which the results cannot agree both with the predictions of quantum mechanics and with the inequality. Extensions of Bell's original theorem made such experimental tests feasible, and more than a dozen of these have been performed, with results overwhelmingly supporting quantum mechanics.[3]

Recently, Greenberger, Horne, and Zeilinger (GHZ)[4] have demonstrated Bell's theorem in a new way, by analyzing a system consisting of three or more correlated spin-1/2 particles. Unlike Bell's original theorem and variants of it, GHZ's demonstration of the incompatibility of quantum mechanics with EPR's propositions concerns only perfect correlations rather than statistical correlations, and it completely dispenses with inequalities. Since EPR's argument for the incompleteness of quantum mechanics was based upon perfect correlations, GHZ's analysis lies close to the heart of EPR's ideas, but with the surprising turnabout of exhibiting a contradiction. GHZ's demonstration can also be read as a gedankenexperiment for testing quantum mechanical predictions against the propositions of EPR on locality, reality, and completeness, and it may be possible, as we shall discuss, to transform the gedankenexperiment into a real experiment. There was one previous proof of Bell's theorem that dispensed with inequalities, that of Kochen and (equivalently) of Heywood and Redhead,[5] who analyzed a pair of spin-1 systems in a state of total spin angular momentum zero. GHZ's demonstration has several advantages, however. First, it is much shorter, as a result of the freedom for manipulation afforded by three or more particles. Second, GHZ's argument suggests a very direct gedankenexperiment. And third, if GHZ's gedankenexperiment can be realized, then a new type of multiparticle correlation experiment will be initiated.

The central purpose of this paper is to present and develop GHZ's original argument, which has been published so far only in outline in proceedings of a conference. As background we shall recapitulate in Sec. II the route from EPR's propositions to Bell's theorem of 1964, via the gedankenexperiment of Bohm,[6] who introduced the use of a pair of spin-1/2 particles. We emphasize that the premisses of Bell are those of EPR (adapted to Bohm's gedankenexperiment), but that he revealed an unexpected consequence of their premisses. In Sec. III we shall present GHZ's argument, which uses a gedankenexperiment with three or four correlated spin-1/2 particles in order to prove Bell's theorem without resorting to an inequality. Section IV will present a new gedankenexperiment involving three particles, as in versions of GHZ's proof by Greenberger and Choi[7] and by Mermin.[8] The innovation is that in the new experiment propagation directions rather than spins are correlated, as in recent two-particle interferometer experiments.[9] In Sec. V we shall show how EPR's program might be feasibly tested experimentally, by observing approximations to theoretically perfect correlations. Section VI will discuss the possibility of real multiparticle experiments, by generating triples and quadruples of photons, either via atomic cascades or via down conversion.

It is our intention to make this paper as nearly self-contained as possible. We shall present the arguments of EPR, Bell, and GHZ essentially in their entirety, and the new material (of Secs. IV, V, and VI) will not presuppose any previous acquaintance with two-particle interferometry. A series of Appendices will present all the relevant quantum mechanical calculations. We hope, therefore, to make accessible both some important earlier contributions to the foundations of quantum mechanics and some fascinating recent developments.

II. FROM EPR TO BELL'S THEOREM

In this section we shall consider the system represented in Fig. 1, consisting of a pair of spin-1/2 particles produced at a source and moving freely in opposite directions. Particle 1 is subjected to a spin measurement by a Stern–Gerlach apparatus with magnetic field in the $\hat{\mathbf{n}}_1$ direction. The outcome of the measurement will be labeled $+1$ if the $\hat{\mathbf{n}}_1$ component of spin is found to be up, and -1 if down. Particle 2 is similarly subjected to a measurement by Stern–Gerlach apparatus with magnetic field along $\hat{\mathbf{n}}_2$. We shall assume that the pair is produced with total spin angular momentum zero. Then the quantum state is

$$|\Psi\rangle = (1/\sqrt{2})[|+\rangle_1|-\rangle_2 - |-\rangle_1|+\rangle_2], \tag{1}$$

where the kets $|+\rangle_1$ and $|-\rangle_1$ represent states of spin-up and -down, respectively, along the arbitrary direction $\hat{\mathbf{n}}$ for particle 1, and $|+\rangle_2$ and $|-\rangle_2$ have analogous meanings for particle 2 with the same direction $\hat{\mathbf{n}}$. Quantum mechanics of the spin-1/2 system yields the remarkable result that the state of Eq. (1) is the same for all unit vectors $\hat{\mathbf{n}}$, which is an invariance expressing the spherical symmetry of the state of total spin angular momentum zero, and therefore no ambiguity results from our omission of $\hat{\mathbf{n}}$ from the notation. (See Appendix A.) Another remarkable feature of $|\Psi\rangle$ is that, in Schrödinger's terminology, it is entangled—that is, it cannot be written in any way as a product of single-particle states (Appendix A). The feature of greatest importance for our purposes is that $|\Psi\rangle$ entails perfect spin correlation: If the $\hat{\mathbf{n}}$ component of spin is found to be $+1$ for particle 1, then with certainty it will be found to be -1 for particle 2, and vice versa. (Sometimes this relation is called perfect "anticorrelation," in contrast to the situation in which the spin components of the two particles have the same value, but our simpler terminology should cause no confusion.)

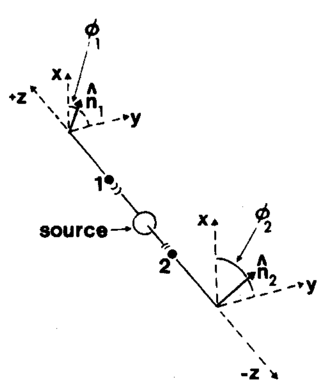

Fig. 1. The Bohm gedankenexperiment. The source emits a pair of spin-1/2 particles, 1 and 2, in the state of Eq. (1). Particle i ($i = 1,2$) enters its own Stern–Gerlach apparatus oriented along direction $\hat{\mathbf{n}}_i$. Behind each apparatus two detectors, not shown, record whether the result is up or down.

From the state $|\Psi\rangle$ of Eq. (1) one can calculate (see Appendix B) the joint probabilities $P^{\Psi}_{++}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2)$, $P^{\Psi}_{+-}(\mathbf{n}_1,\hat{\mathbf{n}}_2)$, $P^{\Psi}_{-+}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2)$, $P^{\Psi}_{--}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2)$, where the first subscript indicates whether the outcome of the measurement on particle 1 is $+1$ or -1, and the second subscript is analogous for particle 2, and where $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$ are the directions along which the spin is measured. The expectation value of the product of the measurement outcomes is defined as

$$E^{\Psi}(\hat{\mathbf{n}}_1\hat{\mathbf{n}}_2) = P^{\Psi}_{++}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2) - P^{\Psi}_{+-}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2) - P^{\Psi}_{-+}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2) + P^{\Psi}_{--}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2). \tag{2}$$

As one might anticipate from the rotational invariance of the state $|\Psi\rangle$, this expectation value depends only upon the angle between the directions $(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2)$, specifically (Appendix B),

$$E^{\Psi}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2) = -\hat{\mathbf{n}}_1\cdot\hat{\mathbf{n}}_2. \tag{3}$$

In the special cases of $\hat{\mathbf{n}}_1 = \hat{\mathbf{n}}_2$, Eq. (3) expresses the perfect correlation mentioned previously.

So far we have given the quantum mechanical description of a pair of spin-1/2 particles in the quantum state $|\Psi\rangle$. We now present EPR's argument (adapted to Bohm's gedankenexperiment) that this quantum mechanical description of the pair of particles cannot be complete. The first of their premisses is drawn from quantum mechanics, and the other three are quite plausible propositions about locality, reality, and completeness, which we state in EPR's words.

(i) *Perfect correlation*: If the spins of particles 1 and 2 are measured along the same direction, then with certainty the outcomes will be found to be opposite.

(ii) *Locality*: "Since at the time of measurement the two systems no longer interact, no real change can take place in the second system in consequence of anything that may be done to the first system."

(iii) *Reality*: "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."

(iv) *Completeness*: "Every element of the physical reality must have a counterpart in the [complete] physical theory."

EPR's argument now proceeds as follows. Because of perfect correlation (i), we can predict with certainty the result of measuring any component of spin of particle 2 by previously choosing to measure the same component of spin of particle 1. By locality (ii), the measurement performed on particle 1 can cause no real change in particle 2. Hence, by the premiss about reality (iii), the chosen spin component of particle 2 is an element of physical reality. But this argument goes through for any component of spin, and hence all of the components of spin of particle 2 are elements of physical reality[10] (and, of course, the same is true of particle 1, by a parallel argument). There is, however, no quantum state of a spin-1/2 particle in which all components of spin have definite values. Therefore, by (iv), quantum mechanics cannot be a complete theory; at least in the case of a pair of spin-1/2 particles with total spin angular momentum zero, there are elements of physical reality for which quantum mechanics has no counterpart. Having shown from their premisses that quantum mechanics is incomplete, EPR do not exhibit a completion or even a model of a completion. They do say, "We believe ... that such a theory is possible."

Bell's argument of 1964 commences by recapitulating the argument of EPR.[11] He then introduces the notation λ for the complete state of a pair of particles, specifying all of the elements of physical reality of the pair at some suitable instant, and he notes, "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." If λ belongs to the set of complete states consistent with the perfect correlation cases of Eq. (3) (a set which we shall denote by $\Lambda_{\hat{n}}$), then λ predetermines the outcome of a measurement of the $\hat{n}$ component of spin of both particles 1 and 2. Hence, if the spin-up and spin-down outcome channels of the Stern–Gerlach analyzers are labeled $+1$ and -1, then there exist functions $A_\lambda(\hat{n})$ and $B_\lambda(\hat{n})$, with values ± 1, defined for all $\hat{n}$ and all $\lambda \in \Lambda_{\hat{n}}$ which are the outcomes, respectively, for particles 1 and 2.

He then considers a probability measure ρ on the whole space of complete states Λ, in order to give a statistical characterization of the ensemble of pairs prepared in the quantum state $|\Psi\rangle$. EPR's program clearly demands the use of probability, because the diversity of measurement results shows that the individual pairs of the ensemble cannot all be in the same complete state. Expectation values of all physical quantities of interest can be defined in terms of ρ. Specifically, the expectation value of the product of outcomes of spin measurements on the two particles is

$$E^{\rho}(\hat{n}_1,\hat{n}_2) = \int_{\Lambda} A_\lambda(\hat{n}_1) B_\lambda(\hat{n}_2) d\rho. \tag{4}$$

Bell emphasizes a fact that has been built into the formalism: that in this expression for the expectation value, the factor $A_\lambda(\hat{n}_1)$ is independent of $\hat{n}_2$ and the factor $B_\lambda(\hat{n}_2)$ is independent of $\hat{n}_1$—as required by the locality assumption (ii) of EPR. Since EPR's argument for their program commenced with the perfect quantum mechanical correlations [assumption (i)], it is essential that the expectation value of Eq. (4) agree with that of Eq. (3) when $\hat{n}_1 = \hat{n}_2 = \hat{n}$:

$$E^{\rho}(\hat{n},\hat{n}) = E^{\Psi}(\hat{n},\hat{n}) = -1. \tag{5}$$

Equation (5) is a very strong constraint. Since the minimum possible value of $A_\lambda(\hat{n})B_\lambda(\hat{n})$ is -1 (the only other possible value being $+1$), it is impossible to satisfy Eq. (5) unless the set of λ's for which $A_\lambda(\hat{n}) = -B_\lambda(\hat{n})$ has probability measure unity, i.e., in the notation introduced above, unless $\rho(\Lambda_{\hat{n}})$ is unity for each $\hat{n}$. It is worthwhile to point out that there is no ground in EPR's argument for concluding that the $\Lambda_{\hat{n}}$'s for different $\hat{n}$ are identical, nor is this identity needed for Bell's demonstration.[12]

Bell now has the concepts required for his theorem. The theorem states that the envisaged complete theory (with $A_\lambda(\hat{n}_1), B_\lambda(\hat{n}_2)$, and ρ as described) must disagree with some of the statistical predictions of quantum mechanics. He uses the expression for expectation values given in Eq. (4) and, by a simple mathematical argument (Appendix C), shows that

$$|E^{\rho}(\hat{a},\hat{b}) - E^{\rho}(\hat{a},\hat{c})| - E^{\rho}(\hat{b},\hat{c}) - 1 \leqslant 0. \tag{6}$$

This is known as "Bell's inequality": Or, rather, it is the first of a family of inequalities that have collectively been given that name.

The remainder of the proof consists in noting that there are choices of directions $\hat{a},\hat{b},\hat{c}$ for which the quantum mechanical expectation values of Eq. (3) conflict with inequality (6). For simplicity, rewrite Eq. (3) for the special case in which both $\hat{n}_1$ and $\hat{n}_2$ lie in the x-y plane, so that they are identified by their azimuthal angles ϕ_1 and ϕ_2, and then

$$E^{\Psi}(\phi_1,\phi_2) = -\cos(\phi_1 - \phi_2). \tag{3'}$$

If $\hat{a},\hat{b},\hat{c}$ lie in the x-y plane with azimuthal angles 0, $\pi/3$, and $2\pi/3$, respectively, the discrepancy between Bell's inequality and quantum mechanics emerges:

$$E^{\Psi}(\hat{a},\hat{b}) = E^{\Psi}(\hat{b},\hat{c}) = -\tfrac{1}{2}, \quad E^{\Psi}(\hat{a},\hat{c}) = +\tfrac{1}{2}, \tag{3''}$$

so that

$$|E^{\Psi}(\hat{a},\hat{b}) - E^{\Psi}(\hat{a},\hat{c})| - E^{\Psi}(\hat{b},\hat{c}) - 1 = \tfrac{1}{2}, \tag{6'}$$

in disagreement with inequality (6). Hence, no choice of the λ's, the functions A and B, and the probability measure ρ on the space of complete states can yield agreement with the quantum mechanical predictions of Eq. (3), if these choices conform to premisses (i) through (iv). This is Bell's theorem of 1964.

Before proceeding with the new multiparticle discussion of Sec. III, we emphasize several important differences between it and Bell's two-particle theorem. First, the perfect correlation (i) and the other EPR premisses (ii)–(iv) are self-consistent in the case of two spin-1/2 particles (see Appendix D). However, as we shall see in Sec. III, the perfect correlations in a three- (or more) particle system are not consistent with the other EPR premisses. Thus, for such a system, the EPR program cannot be carried out even for the special case of perfect correlations.

Second, in the system of two spin-1/2 particles, contradictions develop only when one considers the quantum mechanical statistical predictions. This incompatibility is demonstrated by deriving an inequality from EPR's premisses and then noting that the quantum mechanical statistical predictions do not satisfy this inequality. However, in the three-particle system, there is no point in deriving an inequality, or anything else for that matter, since the premisses are inconsistent.

Third, in the case of two spin-1/2 particles with total spin zero, the cosine of Eq. (3′) plays a central role in proving that quantum mechanics contradicts the inequality. However, in the three-particle case, the specific form of the correlation plays no role in demonstrating a contradiction.

III. BELL'S THEOREM WITHOUT INEQUALITIES

Consider a system of four spin-1/2 particles produced so that particles 1 and 2 move freely in the positive z-direction

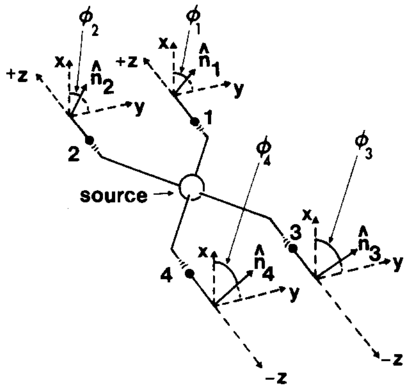

Fig. 2. A four-particle gedankenexperiment. The source emits a quadruple of spin-1/2 particles, 1,2,3, and 4, in the state of Eq. (7). Particle i ($i = 1,2,3,4$) enters its own Stern–Gerlach apparatus oriented along direction $\hat{\mathbf{n}}_i$. We emphasize that the four Stern–Gerlach apparatuses can be separated by arbitrarily large distances. Behind each apparatus two detectors, not shown, record whether the result is up or down.

and particles 3 and 4 in the negative z direction, as shown in Fig. 2. As indicated in the figure, the beams bearing particles 1 and 2 are spatially displaced,[13] so that they can enter different Stern–Gerlach analyzers with orientations $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$, respectively. Similarly, let $\hat{\mathbf{n}}_3$ and $\hat{\mathbf{n}}_4$ be the orientations of Stern–Gerlach analyzers receiving particles 3 and 4. If the four particles result from the decay of a single spin-1 particle into a pair of spin-1 particles, each of which then decays into a pair of spin-1/2 particles, with the z component of spin initially zero and remaining so throughout the decay process, then the quantum mechanical spin state of the four particles is

$$|\Psi\rangle = (1/\sqrt{2})[|+\rangle_1|+\rangle_2|-\rangle_3|-\rangle_4 - |-\rangle_1|-\rangle_2|+\rangle_3|+\rangle_4] \qquad (7)$$

(see Appendix E). Consider, as in Sec. II, the expectation value $E^{\Psi}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2,\hat{\mathbf{n}}_3,\hat{\mathbf{n}}_4)$ of the product of the outcomes when the orientations are as indicated. It is shown in Appendix F that

$$E^{\Psi}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2,\hat{\mathbf{n}}_3,\hat{\mathbf{n}}_4) = \cos\theta_1 \cos\theta_2 \cos\theta_3 \cos\theta_4 - \sin\theta_1 \sin\theta_2 \sin\theta_3 \sin\theta_4 \times\cos(\phi_1+\phi_2-\phi_3-\phi_4), \qquad (8)$$

where (θ_1,ϕ_1) are the polar and azimuthal angles of $\hat{\mathbf{n}}_1$, etc.

For simplicity, we shall restrict our attention to $\hat{\mathbf{n}}$'s in the x-y plane, so that

$$E^{\Psi}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2,\hat{\mathbf{n}}_3,\hat{\mathbf{n}}_4) = -\cos(\phi_1+\phi_2-\phi_3-\phi_4). \qquad (9)$$

Of particular interest will be the following cases of perfect correlation:

$$\text{If } \phi_1+\phi_2-\phi_3-\phi_4 = 0, \text{ then } E^{\Psi}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2,\hat{\mathbf{n}}_3,\hat{\mathbf{n}}_4), = -1, \qquad (10a)$$

$$\text{If } \phi_1+\phi_2-\phi_3-\phi_4 = \pi, \text{ then } E^{\Psi}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2,\hat{\mathbf{n}}_3,\hat{\mathbf{n}}_4) = +1. \qquad (10b)$$

EPR's premisses can be adapted to the present four-particle situation as follows.

(i) *Perfect correlation*: With four Stern–Gerlach analyzers set at angles satisfying the conditions of either (10a) or (10b), knowledge of the outcomes for any three particles enables a prediction with certainty of the outcome for the fourth.

(ii) *Locality*: Since at the time of measurement the four particles are arbitrarily far apart, they presumably do not interact, and hence no real change can take place in any one of them in consequence of what is done to the other three.

(iii) *Reality*: same as in Sec. II.

(iv) *Completeness*: Same as in Sec. II.

It should be emphasized that (i) has a very different status from that of (ii)–(iv). Premiss (i) is a prediction of the quantum mechanical state of Eq. (7) in the arrangement of the gedankenexperiment of Fig. 2, whereas premisses (ii)–(iv) are propositions in EPR's world view—propositions that are quite plausible, however, and which agree well with classical physics. The premisses are listed sequentially, in spite of the difference of their status, because they play crucial roles in EPR's argument for their program.

The arguments in Sec. II can now be paralleled to infer the existence of four functions $A_\lambda(\phi_1)$, $B_\lambda(\phi_2)$, $C_\lambda(\phi_3)$, $D_\lambda(\phi_4)$ with the values $+1$ or -1. These functions are the outcomes of spin measurements along the respective directions when the complete state of the four particles is λ. One could now proceed to derive inequalities of Bell's type and to reveal discrepancies with the statistical predictions of Eq. (9) for appropriate choices of $\phi_1,\phi_2,\phi_3,\phi_4$. There is no point in doing so, however, until the consistency of premisses (i)–(iv) is established. One might suspect that this could easily be done, because in the case of a pair of spin-1/2 particles there already exists a model of Bell (Appendix D) exhibiting the consistency of (ii)–(iv) with the quantum mechanical perfect correlation. That is, one might expect to construct a model, parallel to Bell's, which explicitly exhibits functions A, B, C, and D reproducing the perfect correlations of (10a) and (10b). However, this expectation will not be fulfilled. No such model is possible. As we shall now proceed to show, EPR's premisses for the four-particle situation are inconsistent.

First, we restate (10a) and (10b) in terms of the functions A, B, C, and D, the existence of which follows from the premisses:[14]

$$\text{If } \phi_1+\phi_2-\phi_3-\phi_4 = 0, \text{ then } A_\lambda(\phi_1)B_\lambda(\phi_2)C_\lambda(\phi_3)D_\lambda(\phi_4) = -1, \qquad (11a)$$

$$\text{If } \phi_1+\phi_2-\phi_3-\phi_4 = \pi, \text{ then } A_\lambda(\phi_1)B_\lambda(\phi_2)C_\lambda(\phi_3)D_\lambda(\phi_4) = 1. \qquad (11b)$$

Let us now consider some implications of just one of (11a) and (11b), say, the first. Four instances of (11a) are

$$A_\lambda(0)B_\lambda(0)C_\lambda(0)D_\lambda(0) = -1, \qquad (12a)$$

$$A_\lambda(\phi)B_\lambda(0)C_\lambda(\phi)D_\lambda(0) = -1, \qquad (12b)$$

$$A_\lambda(\phi)B_\lambda(0)C_\lambda(0)D_\lambda(\phi) = -1, \qquad (12c)$$

$$A_\lambda(2\phi)B_\lambda(0)C_\lambda(\phi)D_\lambda(\phi) = -1. \qquad (12d)$$

From Eqs. (12a) and (12b) we obtain

$$A_\lambda(\phi)C_\lambda(\phi) = A_\lambda(0)C_\lambda(0), \qquad (13a)$$

and from Eqs. (12a) and (12c) we obtain

$$A_\lambda(\phi)D_\lambda(\phi) = A_\lambda(0)D_\lambda(0). \qquad (13b)$$

A consequence of these is

$$C_\lambda(\phi)/D_\lambda(\phi) = C_\lambda(0)/D_\lambda(0), \qquad (14a)$$

which can be rewritten as

$$C_\lambda(\phi)D_\lambda(\phi) = C_\lambda(0)D_\lambda(0), \quad (14b)$$

because $D_\lambda(\phi)$ is ± 1 and hence equals its inverse, and the same for $D_\lambda(0)$. We then obtain from Eqs. (12d) and (14b)

$$A_\lambda(2\phi)B_\lambda(0)C_\lambda(0)D_\lambda(0) = -1, \quad (15)$$

which in combination with Eq. (12a) yields[15]

$$A_\lambda(2\phi) = A_\lambda(0) = \text{const for all } \phi. \quad (16)$$

Equation (16) is a quite surprising preliminary result. By itself, this equation is not mathematically contradictory, but physically it is very troublesome: For if $A_\lambda(\phi)$ is intended, as EPR's program suggests, to represent an intrinsic spin quantity, then $A_\lambda(0)$ and $A_\lambda(\pi)$ would be expected to have opposite signs. The trouble becomes manifest, and an actual contradiction emerges, when we use (11b)—which until now has not been brought into play—to obtain

$$A_\lambda(\theta+\pi)B_\lambda(0)C_\lambda(\theta)D_\lambda(0) = 1, \quad (17)$$

which in combination with Eq. (12b) yields

$$A_\lambda(\theta+\pi) = -A_\lambda(\theta). \quad (18)$$

This result *confirms* the sign change that we anticipated on physical grounds in EPR's program, but it also *contradicts* the earlier result of Eq. (16) (let $\phi = \pi/2, \theta = 0$). We have thus brought to the surface an inconsistency hidden in premisses (i)–(iv).

In the foregoing algebra, the argument of the function $B_\lambda(\phi_2)$ was fixed throughout to be 0, which shows that premisses (i)–(iv) are also inconsistent when applied to a system of three spin-1/2 particles. But we know from Bell's model (Appendix D) that the corresponding premisses are consistent for a pair of spin-1/2 particles. Correlations of three or more spin-1/2 particles involve at least one more degree of freedom than one finds in correlations of two spin-1/2 particles, and it is clear that the manipulation of an additional degree of freedom is essential to the exhibition of a contradiction.[16] The most significant feature of the new argument is the revelation that the EPR program cannot handle even the perfect correlations of quantum mechanics for systems of three or more particles. There is an irony in this result in that perfect correlations are central to EPR's argument for the existence of states more complete than those of quantum mechanics.

IV. ... AND WITHOUT SPIN

In this section we shall present a new gedankenexperiment, with three particles and without spin, for which the argument of Sec. III applies unchanged. The new system is an extension of a two-particle interferometer previously considered by Horne and Zeilinger[17] for exhibiting entangled states in momentum and position, and it resembles Mermin's[8] figure of a three-spin "gadget." Since the gedankenexperiment does not involve spin, it emphasizes, yet again, that Bell's theorem does not hinge on spin.[17] Since it employs only three particles, it emphasizes that the GHZ argument goes through for three particles. Finally, the new gedankenexperiment may be realizable in the laboratory.

Consider a particle that can decay into three particles of equal mass or into three photons (though, with a slight modification, we could allow any combination of different mass particles and photons). Imagine that the decaying

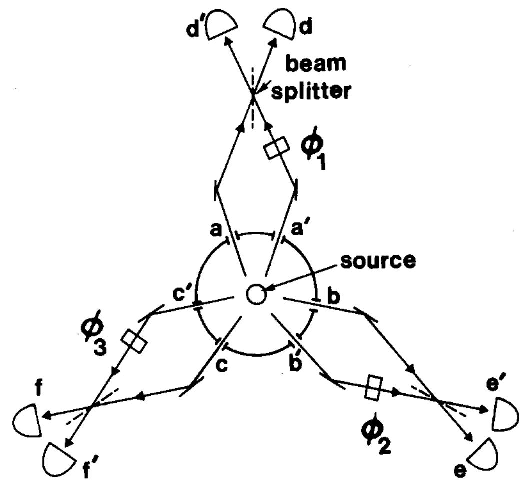

Fig. 3. A gedanken three-particle interferometer. The source emits a triple of particles, 1, 2, and 3, in six beams, with the state given by Eq. (19). A phase shift ϕ_1 is imparted to beam a' of particle 1, and beams a and a' are brought together on a beam splitter before illuminating detectors d and d'. Likewise for particles 2 and 3.

particle has (mean) momentum zero and that the decay occurs in the central region of the arrangement of Fig. 3. If all three particles have the same energy, then, by momentum conservation, they must be emitted 120° apart. The equal energy requirement can be enforced by placing energy filters at the detectors. The central source is surrounded by an array of six apertures: a, b, and c at 120° separation, and a', b', and c' also at 120° separation. Because of the placement of the apertures, the three particles 1, 2, and 3 must emerge either through a, b, and c or through a', b', and c'. Thus the state of the three particles beyond the apertures will be the superposition

$$|\Psi\rangle = (1/\sqrt{2})[|a\rangle_1|b\rangle_2|c\rangle_3 + |a'\rangle_1|b'\rangle_2|c'\rangle_3], \quad (19)$$

where $|a\rangle_1$ denotes the particle 1 in beam a, etc. The insertion of an arbitrary (but fixed) phase between the two terms of Eq. (19) produces no significant change of predictions.

Beyond the apertures beams a and a' are totally reflected so as to overlap at a 50–50 beam splitter,[18] and the two outgoing beams are monitored by detectors d and d'. Note that en route beam a' passes through a phase plate[18] which causes an adjustable phase shift ϕ_1. Consequently, the evolution of the kets $|a\rangle_1$ and $|a'\rangle_1$ is given by

$$|a\rangle_1 \to (1/\sqrt{2})[|d\rangle_1 + i|d'\rangle_1] \quad (20a)$$

and

$$|a'\rangle_1 \to (1/\sqrt{2})e^{i\phi_1}[|d'\rangle_1 + i|d\rangle_1], \quad (20b)$$

where the ket $|d\rangle_1$ denotes particle 1 directed toward detector d, etc. In (20a) and (20b) the i is due to reflection from the beam splitter,[19] and the ϕ_1 is the phase shift due to the phase plate inserted in beam a'. The particle 2 beams and the particle 3 beams are subjected to similar treatment and hence undergo similar evolutions. When the evolutions of (20a) and (20b) and the similar ones for particles 2 and 3 are combined with the initial three-particle state of Eq. (19), a state with eight terms develops (Appendix G), from which we obtain amplitudes and hence probabilities of detection of the three particles by the triple of detectors (d,e,f), the triple of detectors (d',e,f), etc. We assume that the detectors are perfect, so that every triple of particles causes either d or d' to fire, either e or e' to fire, and either f or f' to fire.

The resulting expressions for the probabilities of detections (Appendix G) are

$$P^{\Psi}_{def}(\phi_1,\phi_2,\phi_3) = \tfrac{1}{8}[1 + \sin(\phi_1 + \phi_2 + \phi_3)], \quad (21a)$$

$$P^{\Psi}_{d'ef}(\phi_1,\phi_2,\phi_3) = \tfrac{1}{8}[1 - \sin(\phi_1 + \phi_2 + \phi_3)], \quad (21b)$$

etc. (If the number of primes on the detector labels is even, there is a plus sign; if odd, there is a minus sign.) Obviously the sum of these probabilities for eight possible outcomes (d,e,f), etc. is unity. Paralleling the discussion of Sec. III we call the result $+1$ when a particle enters an unprimed detector (for example, d) and -1 when it enters a primed detector (for example, d'). With these values and the probabilities of Eqs. (21a) and (21b), etc., we may calculate the expectation value of the product of the three outcomes. The result (Appendix G) is

$$E^{\Psi}(\phi_1,\phi_2,\phi_3) = \sin(\phi_1 + \phi_2 + \phi_3). \quad (22)$$

Perfect correlations are obtained for the following choices of angles:

$$\text{If } \phi_1 + \phi_2 + \phi_3 = \pi/2, \text{ then } E^{\Psi}(\phi_1,\phi_2,\phi_3) = +1, \quad (23a)$$

$$\text{If } \phi_1 + \phi_2 + \phi_3 = 3\pi/2, \text{ then } E^{\Psi}(\phi_1,\phi_2,\phi_3) = -1. \quad (23b)$$

All the requirements are now met for the argument of Sec. III to go through, exhibiting a contradiction in EPR's premisses. Note that Eqs. (22), (23a), and (23b) are the analogs of Eqs. (9), and (10a) and (10b), the significant change being that the phases associated with the plates take the place of the azimuthal angles of the Stern–Gerlach orientations.

We conclude this section by emphasizing some fascinating features and generalization of the arrangement of Fig. 3, which we call a three-particle interferometer. First, if detectors d,e,f (for example) were monitored for three-particle coincidences, then Eq. (21a) predicts that the observed count rate for these coincidences will depend sensitively on the phases imparted by the phase plates. For example, if $\phi_1 + \phi_2 + \phi_3$ is varied linearly in time (by manipulating one or more of the plates), then the three-particle coincidence rate will vary sinusoidally, with a minimum of zero and a maximum of one-quarter the rate at which triples of particles are emitted from the apertures. We shall call these sinusoidal oscillations three-particle interference fringes. Second, the three-particle interferometer will not exhibit any two-particle fringes. That is, if detectors e and f (for example) are monitored for two-particle coincidences, while detectors d and d' are ignored, the observed two-particle coincidence rate will be completely independent of the phases. This statement follows from the fact that the sum of Eqs. (21a) and (21b) is $\frac{1}{4}$, independent of $\phi_1 + \phi_2 + \phi_3$. Third, the three-particle interferometer will also exhibit no single-particle fringes. If, for example, detector f is monitored while d, d', e, and e' are ignored, then the observed count rate will be one-half the rate of emission of triples through the apertures, independent of the phases. This follows by summing $P^{\Psi}_{def}(\phi_1,\phi_2,\phi_3)$, $P^{\Psi}_{d'ef}(\phi_1,\phi_2,\phi_3)$, $P^{\Psi}_{de'f}(\phi_1,\phi_2,\phi_3)$, and $P^{\Psi}_{d'e'f}(\phi_1,\phi_2,\phi_3)$. Finally, the three-particle interferometer of Fig. 3 and the initial three-particle state of Eq. (19) can be generalized to n-particle interferometry, and the features just described also generalize: An n-particle interferometer will exhibit n-particle fringes, but it will not exhibit $n-1$, $n-2$,..., or single-particle fringes.[20]

V. ON TESTING EPR'S PROGRAM

The new demonstration of Bell's theorem without inequalities does not necessarily call for an experiment, any more than did Bell's original theorem of 1964. However, EPR's program is plausible, and one may suspect that in a situation of conflict between that program and some predictions of quantum mechanics, the latter will turn out to be false. Indeed, this motivation led to more than a dozen experimental tests in the last two decades, with results overwhelmingly supporting quantum mechanics.[3] Is there any point in designing yet another experiment along new lines, in order to rerefute the program of EPR? We think so, for two reasons. The first is sheer intellectual challenge. We would like to know what experiment would have been appropriate had history been different and had GHZ's demonstration been the first proof of Bell's theorem. The second is that the investigation of correlations among three or more particles can open a new, beautiful, and fruitful type of experimentation, of interest independently of EPR. Since multiparticle interferometry seems particularly promising, we shall focus attention on the arrangement of Fig. 3.

It must be emphasized that the experiment that we shall propose will test EPR's program—by which we mean (in the context of Fig. 3) the existence of the complete states λ, the result functions $A_\lambda(\phi_1)$, $B_\lambda(\phi_2)$, $C_\lambda(\phi_3)$, and probability measure over the space of complete states—but will not test the premisses from which EPR argue for their program. The reason for targeting the experiment in this way is that premisses (i)–(iv) together are inconsistent, as we have already seen, while premisses (ii)–(iv) without (i) do not suffice to make testable predictions. We are struck by the curiosity of the logical situation: Perfect correlations are needed to initiate the EPR argument, and perfect correlations suffice to show that the EPR premisses are invalid. By focusing on the pıogram, rather than the argument for the program, we circumvent the difficulty posed by this logical situation. Furthermore, there are good historical and philosophical reasons for considering EPR's program as a hypothesis. Einstein urged his program as early as the Fifth Solvay Congress of 1927, because of his belief that a fundamental physical theory should not be stochastic. The EPR argument of 1935 was intended to support Einstein's program and his interpretation of quantum mechanics without relying upon an antecedent commitment to determinism, but the program itself is worthy of serious consideration even without their argument.

In proposing an experiment we proceed in three steps. First, we show that the mere existence of the result functions $A_\lambda, B_\lambda, C_\lambda$ imposes a remarkably strong constraint on the probability measure ρ [see inequality (29)]. This inequality is in principle testable. The second step is to show how the test could be done even with low-efficiency detectors, provided that we make a plausible auxiliary assumption, which we call fair sampling. Finally, we show that the auxiliary assumption is dispensable if detector efficiencies exceed 90.8%.

In the arrangement of Fig. 3 consider four different choices of the phase angles (ϕ_1,ϕ_2,ϕ_3): namely, $(\pi/2,0,0)$, $(0,\pi/2,0)$, $(0,0,\pi/2)$, and $(\pi/2, \pi/2, \pi/2)$. Each triple of particles emitted through the apertures is assumed by EPR's program to be in a complete state λ, and the assumed existence of the result functions $A_\lambda(\phi_1), B_\lambda(\phi_2), C_\lambda(\phi_3)$ implies λ will predetermine the outcome for each particle for each of the four choices of phase angles. Of course, any specific triple of particles can be subjected to only *one* of the four choices of phase angles, and therefore the entire ensemble of triples emitted from the aperture is subdivided into four mutually exclusive and exhaustive subensembles. But since the triples are emitted before encountering the phase plates, where the subdivision into subensembles takes place, it is reasonable to assume that the same probability measure ρ governs all four of the subensembles.

Consider now the following three statements:

$$A_\lambda(\pi/2)B_\lambda(0)C_\lambda(0) = 1, \qquad (24a)$$

$$A_\lambda(0)B_\lambda(\pi/2)C_\lambda(0) = 1, \qquad (24b)$$

$$A_\lambda(0)B_\lambda(0)C_\lambda(\pi/2) = 1. \qquad (24c)$$

Multiplying Eqs. (24a), (24b), and (24c), as suggested by Mermin,[15] we obtain

$$A_\lambda(\pi/2)B_\lambda(\pi/2)C_\lambda(\pi/2) = 1, \qquad (24d)$$

since the other factors in the product obviously multiply to unity. Consequently, the statement

$$A_\lambda(\pi/2)B_\lambda(\pi/2)C_\lambda(\pi/2) = -1 \qquad (24e)$$

implies that at least one of Eqs. (24a), (24b), and (24c) is false. We can express this last implication in the language of set theory:

$$\Lambda_4 \subseteq \overline{\Lambda}_1 \cup \overline{\Lambda}_2 \cup \overline{\Lambda}_3. \qquad (25)$$

Here,

Λ_1 = the set of all λ's such that Eq. (24a) holds, (26a)

Λ_2 = the set of all λ's such that Eq. (24b) holds, (26b)

Λ_3 = the set of all λ's such that Eq. (24c) holds, (26c)

but (note well)

Λ_4 = the set of all λ's such that Eq. (24e) holds, (26d)

and $\overline{\Lambda}_1$ is the complement of Λ_1 (i.e., the set of all λ's not belonging to Λ_1), etc.

We are now in a position to make use of the probability measure ρ on the space of complete states. Specifically, by (25)

$$\rho(\Lambda_4) \subseteq \leqslant \rho(\overline{\Lambda}_1 \cup \overline{\Lambda}_2 \cup \overline{\Lambda}_3). \qquad (27)$$

But by standard probability theory

$$\rho(\overline{\Lambda}_1 \cup \overline{\Lambda}_2 \cup \overline{\Lambda}_3) \leqslant \rho(\overline{\Lambda}_1) + \rho(\overline{\Lambda}_2) + \rho(\overline{\Lambda}_3), \qquad (28)$$

the reason for connecting the left-hand side to the right-hand side by $\leqslant$ being the possibility that the three sets $\overline{\Lambda}_1$, $\overline{\Lambda}_2$, $\overline{\Lambda}_3$, are not mutually exclusive. From inequalities (27) and (28) we obtain the strong constraint on the probability measure ρ which we have been seeking:

$$\rho(\Lambda_4) \leqslant \rho(\overline{\Lambda}_1) + \rho(\overline{\Lambda}_2) + \rho(\overline{\Lambda}_3). \qquad (29)$$

Inequality (29) is testable, once appropriate connection is supplied between the probabilities $\rho(\Lambda_4)$, etc. and laboratory quantities, a matter that we are about to discuss. Before doing so, however, we wish to stress that inequality (29) was derived from apparently innocent premisses: nothing but the existence of the complete states λ, the result functions $A_\lambda, B_\lambda, C_\lambda$, and the probability measure ρ. The fact that inequality (29) is nontrivial shows that the premisses are not innocent after all. Especially, they embody EPR's commitment to locality, because the result function A_λ for particle 1 depends only on the phase angle ϕ_1 of the phase plate that it encounters, and not on ϕ_2 and ϕ_3, and likewise B_λ does not depend on ϕ_1 and ϕ_3, and C_λ does not depend on ϕ_1 and ϕ_2.

A test of inequality (29) requires a connection between the quantities entering in this inequality and laboratory data. As pointed out in Sec. IV, the arrangement of Fig. 3 defines an ensemble of triples of particles: namely, those emitted through the apertures into the six beams a,b,c,a',b',c'. Because of imperfection of actual detectors, and the possible loss at actual mirrors and beam splitters, not all of these triplet will be detected. If the detectors are made as nearly alike as possible, then it is reasonable to assume "fair sampling": that is, the number of triples actually detected by any three designated detectors, say d,e,f, with any specified phase angles ϕ_1,ϕ_2,ϕ_3 is proportional to the number that would be detected if the mirrors, beam splitters, and detectors were perfect—with the same constant of proportionality in all cases. This fair sampling assumption was previously proposed by Clauser *et al.*[21] in order to permit a test of one of Bell's inequalities. Given the fair sampling assumption, the probability $\rho(\Lambda_1)$ is determined by making the phase angles (ϕ_1,ϕ_2,ϕ_3) be $(\pi/2,0,0)$ and dividing the number of coincidence counts for which the product of outcomes is $+1$ (i.e., those in which an even number of primed detectors are triggered) by the total number of triples detected. If the choices (ϕ_1,ϕ_2,ϕ_3) are set to be $(0,\pi/2,0)$ and $(0,0,\pi/2)$, then $\rho(\Lambda_2)$ and $\rho(\Lambda_3)$ are likewise determined. In order to determine $\rho(\Lambda_4)$ the phase angles are set at $(\pi/2,\pi/2,\pi/2)$, and the number of counts for which the product of the outcomes is -1 (i.e., an odd number of primed detectors are triggered) is divided by the total number of detectors triggered. Since

$$\rho(\overline{\Lambda}_1) = 1 - \rho(\Lambda_1), \qquad (30)$$

etc., we would then have all of the information in hand to test inequality (29). Violation of the inequality would constitute strong evidence against EPR's program, and of course this is the result that we anticipate, in view of the general success of quantum mechanics.

There has been extensive discussion of possible failure of the fair sampling assumption and of ways to dispense with it in tests of Bell's inequality. It has been shown that the assumption would indeed be dispensable if the detectors were at least 82.8% efficient.[22] A similar conclusion can be drawn in the present context, concerning a test of inequality (29). Suppose the rate of triples emitted through the apertures is known (possibly by some calorimetric method) and for $(\phi_1,\phi_2,\phi_3) = (\pi/2,0,0)$ a fraction f of these is detected with product of outcomes equal to $+1$. Others may be detected with product of outcomes -1, and others are not detected at all. From these data, together with the definition of $\rho(\Lambda_1)$, we obtain

$$\rho(\Lambda_1) \geqslant f \qquad (31a)$$

and

$$\rho(\overline{\Lambda}_1) \leqslant 1 - f. \qquad (31b)$$

Similar bounds may be obtained for $\rho(\Lambda_2)$, $\rho(\overline{\Lambda}_2)$, etc., and for simplicity we shall take the fractions detected with the outcomes mentioned in the definitions of these sets to be the same f. Combining these bounds with inequality

(29) yields

$$f \leqslant \rho(\Lambda_4) \leqslant \rho(\overline{\Lambda}_1) + \rho(\overline{\Lambda}_2) + \rho(\overline{\Lambda}_3) \leqslant 3(1-f). \quad (32)$$

This inequality will be violated (with consequent disconfirmation of EPR's program) if $f > 0.75$. As pointed out above, there are two contributions to the fraction $1-f$: one from the detection of triples in the "wrong" way (i.e., contrary to the prediction of quantum mechanics) and one from the nondetection of triples. The latter part could be as large as $1-\eta^3$, where η is the efficiency of a single detector. Hence, even if no "wrong" triples are detected, there would be certainty that f is greater than 0.75 only if η^3 is greater than 0.75, i.e., only if η is greater than 0.908. Thus the demand on detector efficiency if the fair sampling assumption is to be avoided in the present arrangement is more stringent than in tests of Bell's inequality.[23]

An advocate of EPR's point of view might attempt to salvage part of their program, even if inequality (29) turns out to be refuted experimentally. The salvaging strategy could consist in abandoning the functions $A_\lambda(\phi_1), B_\lambda(\phi_2), C_\lambda(\phi_3)$, and thereby giving up the idea that the complete state λ predetermines the outcomes of measurements. Instead, it would be assumed that when λ and ϕ_1 are given, there is only a definite probability that particle 1 will trigger detector d and a definite probability that it will trigger detector d': and likewise for particles 2 and 3. The spirit of EPR's program would be preserved by making the probabilities concerning one particle depend only on the phase plate that it encounters, and not on the other two phase plates. The type of theory envisaged by this strategy is often called a stochastic local theory, in contrast to a deterministic local theory, which has been considered so far in this paper.

We already know from traditional work on Bell's theorem that this salvaging strategy fails. In the first place, it was shown by Bell[24] in 1971 and by Clauser and Horne[25] in 1974 that the same inequality governing statistical correlations that is derived from a deterministic local theory can be derived from a stochastic local theory. Consequently, the violation of Bell's inequality by the data of numerous experiments constitutes a disconfirmation of the weakened version of EPR's program. Second, there are "equivalence theorems" (Stapp[26] and Fine[27]) to the effect that the predictions of any stochastic local theory can be duplicated by an appropriate deterministic local theory. Hence, any experimental evidence against the family of deterministic local theories would automatically be evidence against the family of stochastic local theories. The second of these two reasons for the failure of the salvaging strategy is the one more relevant to the argument of the present paper, since this reason makes no reference to Bell's inequality.

VI. REAL EXPERIMENTS

There are two immediate avenues for experimental verification of the remarkable features of multiparticle correlations; both are generalizations of previous experiments with pairs of photons. As the first possibility one may contemplate the polarization correlations among three or more photons emitted by a cascading atom. Such an experiment would be a generalization of the many two-photon polarization correlation experiments done over the past two decades, all of which are descendants of the pioneering experiments of Kocher and Commins[28] and of Freedman and Clauser.[3] The second possibility would be to exploit the momentum and energy correlations among three or more photons emitted in the process of parametric down conversion. Such an experiment would generalize the recent series two-photon interference experiments which are all descendants of the pioneering experiments of Burnham and Weinberg[29] and of Gosh and Mandel.[9]

A generalization of the atomic cascade two-photon experiments would use an atom cascading through two intermediate levels to produce three photons. For example, consider an atom undergoing electric dipole transitions that carry the atom from a state of zero total angular momentum through two intermediate states of angular momentum one back to a state of zero total angular momentum, that is a $J=0 \to J=1 \to J=1 \to J=0$ cascade. Suppose for simplicity that the three detectors select three photons that propagate in a plane along directions at angles 120° from each other. Then by conservation of angular momentum the three-photon state is[30]

$$|\Psi\rangle = (1/\sqrt{2})[|R\rangle_1 |R\rangle_2 |R\rangle_3 + |L\rangle_1 |L\rangle_2 |L\rangle_3], \quad (33)$$

which has the required entanglement to exhibit polarization correlations analogous to the direction correlations of Sec. IV.

A generalization of the class of recent two-photon down-conversion experiments would utilize the fact that, at least in principle, down conversion can produce three or more correlated photons. As a specific example consider the three-photon generalization of two-particle interferometry (Horne, Shimony, and Zeilinger[9] and Rarity and Tapster[17]). In such a generalization six apertures a,b,c and a',b',c' would be suitably placed downstream from a down-conversion crystal so that by energy and momentum conservation the emerging state of the three-photon radiation would be

$$|\Psi\rangle = (1/\sqrt{2})[|a\rangle_1 |b\rangle_2 |c\rangle_3 + |a'\rangle_1 |b'\rangle_2 |c'\rangle_3], \quad (34)$$

which is formally the same as Eq. (19) of Sec. IV.

One could equally well contemplate generalization of some of the other existing two-photon down-conversion experiments. An example would be a three-photon generalization of the experiment proposed by Franson[31] and performed by Ou and Mandel[32] and by Kwiat *et al.*[33] Such experiments exploit correlations in energy/time instead of momentum/direction. Alternatively, instead of direct down-conversion to three or four photons, one could in principle contemplate a two-step cascade of two-photon down conversions to produce a four-photon entangled state.

ACKNOWLEDGMENTS

We wish to thank N. D. Mermin for some valuable suggestions made at the Amherst Workshop on Foundations of Quantum Mechanics, 10–15 June 1990. This work was supported by the National Science Foundation under Grant Nos. DMR-8713559 and INT-8713341. We also wish to thank R. Clifton, J. Butterfield, and M. Redhead. During the early stages of this work D.M.G. was a Fulbright fellow at the Atominstitut, Vienna, Austria, and a recipient of a Humboldt Senior Scientist award at the Max-Planck-Institut für Quantenoptik, Garching, F.R.G.; he also wishes to acknowledge partial support from the CUNY–FRAP program. A.Z. acknowledges partial support from the Austrian Fonds zur Förderung der wissenschaftlichen Forschung, grant No. P6635.

APPENDIX A: PERFECT CORRELATIONS IN THE TWO-PARTICLE STATE $|\Psi\rangle$

In order to show the rotational invariance of the state of Eq. (1) we shall have to express the kets $|\hat{\mathbf{n}},+\rangle$ and $|\hat{\mathbf{n}},-\rangle$, the states of spin-up and -down along the $\hat{\mathbf{n}}$ direction, in terms of $|+\rangle$ and $|-\rangle$, which are states of spin-up and -down along a specified direction, commonly $\hat{\mathbf{z}}$. We shall use the standard Pauli matrices $\boldsymbol{\sigma}_x, \boldsymbol{\sigma}_y, \boldsymbol{\sigma}_z$ and write

$$\boldsymbol{\sigma}\cdot\hat{\mathbf{n}}|\hat{\mathbf{n}},+\rangle = |\hat{\mathbf{n}},+\rangle, \tag{A1a}$$

$$\boldsymbol{\sigma}\cdot\hat{\mathbf{n}}|\hat{\mathbf{n}},-\rangle = -|\hat{\mathbf{n}},-\rangle. \tag{A1b}$$

Each of the Eqs. (A1a) and (A1b) can be written as a pair of coupled linear equations, which can be solved to yield (up to an arbitrary scalar multiple)

$$|\hat{\mathbf{n}},+\rangle = (\cos\theta/2)e^{-i\phi/2}|+\rangle + (\sin\theta/2)e^{i\phi/2}|-\rangle, \tag{A2a}$$

$$|\hat{\mathbf{n}},-\rangle = -(\sin\theta/2)e^{-i\phi/2}|+\rangle + \cos\theta/2)e^{i\phi/2}|-\rangle, \tag{A2b}$$

where θ and ϕ are the polar and azimuthal angles of $\hat{\mathbf{n}}$. Now consider

$$|\Psi(\hat{\mathbf{n}})\rangle = (1/\sqrt{2})[|\hat{\mathbf{n}},+\rangle_1|\hat{\mathbf{n}},-\rangle_2 - |\hat{\mathbf{n}},-\rangle_1|\hat{\mathbf{n}}+\rangle_2], \tag{A3}$$

and insert Eqs. (A2a) and (A2b) to obtain

$$|\Psi(\hat{\mathbf{n}})\rangle = (1/\sqrt{2})[|+\rangle_1|-\rangle_2 - |-\rangle_1|+\rangle_2] = |\Psi\rangle. \tag{A4}$$

Thus rotational invariance is proved, and incidentally we are justified in using the notation of Eq. (1), which refrains from making the basis explicit.

The foregoing argument implies as a corollary that $|\Psi\rangle$ is entangled. The reason is that there are no bases for the spin-1/2 particle other than pairs of kets of the form $|\hat{\mathbf{n}},+\rangle$ and $|\hat{\mathbf{n}},-\rangle$, and we have just seen that in none of these bases is $|\Psi\rangle$ expressed as a product state.

It is instructive, however, to give a more general proof of entanglement. Let $|\Phi\rangle$ be a vector representing a state of a two-particle system with the form

$$|\Phi\rangle = c|u\rangle_1|v\rangle_2 + d|u'\rangle_1|v'\rangle_2, \tag{A5}$$

where $|u\rangle_1$ and $|u'\rangle_1$ are orthonormal kets in the space of states of particle 1 and $|v\rangle_2$ and $|v'\rangle_2$ are orthonormal kets in the space of states of particle 2, and both of the scalars c and d are nonzero. We claim that $|\Phi\rangle$ is entangled. If not, it can be written as

$$|\Phi\rangle = |w\rangle_1|z\rangle_2, \tag{A6}$$

where $|w\rangle_1$ and $|z\rangle_2$ are, respectively, in the spaces of states of particles 1 and 2. General vector space considerations imply that $|w\rangle_1$ can be expressed as a linear superposition of $|u\rangle_1$ and $|u'\rangle_1$, and $|z\rangle_2$ can be expressed as a linear superposition of $|v\rangle_2$ and $|v'\rangle_2$, i.e.,

$$|w\rangle_1 = a|u\rangle_1 + a'|u'\rangle_1, \tag{A7a}$$

$$|z\rangle_2 = b|v\rangle_2 + b'|v'\rangle_2. \tag{A7b}$$

Then

$$|w\rangle_1|z\rangle_1 = ab|u\rangle_1|v\rangle_2 + ab'|u\rangle_1|v'\rangle_2 + a'b|u'\rangle_1|v\rangle_2 + a'b'|u'\rangle_1|v'\rangle_2. \tag{A8}$$

Comparing Eqs. (A5) and (A8) yields

$$ab' = a'b = 0, \tag{A9a}$$

which in turn implies that either

$$c = ab = 0 \tag{A9b}$$

or

$$d = a'b' = 0, \tag{A9c}$$

contrary to assumption.

APPENDIX B: STATISTICAL CORRELATIONS IN THE TWO-PARTICLE STATE $|\Psi\rangle$

Because of Eq. (A4) of Appendix A we can write the state $|\Psi\rangle$ of Eq. (1) in the form

$$|\Psi\rangle = (1/\sqrt{2})(|\hat{\mathbf{n}}_1,+\rangle_1|\hat{\mathbf{n}}_1,-\rangle_2 - |\hat{\mathbf{n}}_1,-\rangle_1|\hat{\mathbf{n}}_1,+\rangle_2), \tag{B1}$$

where $\hat{\mathbf{n}}_1$ is an arbitrary direction. Let $\hat{\mathbf{n}}_2$ be another arbitrary direction, which will be taken to be the polar axis, and let the polar angle of $\hat{\mathbf{n}}_1$ with respect to $\hat{\mathbf{n}}_2$ be θ. By proper choice of the other coordinates the azimuthal angle for $\hat{\mathbf{n}}_1$ will be zero. Hence, adapting Eqs. (A2a) and (A2b), we have

$$|\hat{\mathbf{n}}_1,+\rangle_2 = (\cos\theta/2)|\hat{\mathbf{n}}_2,+\rangle_2 + (\sin\theta/2)|\hat{\mathbf{n}}_2,-\rangle_2, \tag{B2a}$$

$$|\hat{\mathbf{n}}_1,-\rangle_2 = -(\sin\theta/2)|\hat{\mathbf{n}}_2,+\rangle_2 + (\cos\theta/2)|\hat{\mathbf{n}}_2,-\rangle_2. \tag{B2b}$$

Hence,

$$\begin{aligned}|\Psi\rangle = (1/\sqrt{2})[&-(\sin\theta/2)|\hat{\mathbf{n}}_1,+\rangle_1|\hat{\mathbf{n}}_2,+\rangle_2 \\ &+(\cos\theta/2)|\hat{\mathbf{n}}_1,+\rangle_1|\hat{\mathbf{n}}_2,-\rangle_2 \\ &-(\cos\theta/2)|\hat{\mathbf{n}}_1,-\rangle_1|\hat{\mathbf{n}}_2,+\rangle_2 \\ &-(\sin\theta/2)|\hat{\mathbf{n}}_1,-\rangle_1|\hat{\mathbf{n}}_2,-\rangle_2].\end{aligned} \tag{B3}$$

The amplitude for the joint outcome up along $\hat{\mathbf{n}}_1$ for particle 1 and up along $\hat{\mathbf{n}}_2$ for particle 2 is $-(\sin\theta/2)/\sqrt{2}$, and similarly for the three other possible outcomes. Hence,

$$P_{+\Psi+}(\hat{\mathbf{n}}_2,\hat{\mathbf{n}}_2) = \tfrac{1}{2}\sin^2\theta/2, \tag{B4a}$$

$$P_{+\Psi-}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2) = \tfrac{1}{2}\cos^2\theta/2, \tag{B4b}$$

$$P_{-\Psi+}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2) = \tfrac{1}{2}\cos^2\theta/2, \tag{B4c}$$

$$P_{-\Psi-}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2) = \tfrac{1}{2}\sin^2\theta/2. \tag{B4d}$$

Using Eqs. (B4a), (B4b), (B4c), and (B4d) and the definition of $E^{\Psi}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2)$ of Eq. (2), we obtain

$$E^{\Psi}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2) = \sin^2\theta/2 - \cos^2\theta/2 = -\cos\theta = -\hat{\mathbf{n}}_1\cdot\hat{\mathbf{n}}_2. \tag{B5}$$

Of course, this expectation value can also be computed using

$$E^{\Psi}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2) = \langle\Psi|(\hat{\mathbf{n}}_1\cdot\boldsymbol{\sigma}_1)(\hat{\mathbf{n}}_2\cdot\boldsymbol{\sigma}_2)|\Psi\rangle. \tag{B6}$$

The reader should be on guard against confusing the *pure* two-particle entangled state $|\Psi\rangle$ with a rotationally invariant *mixture* of product states, which will not yield correlations as strong as $|\Psi\rangle$ does. Consider specifically the following ensemble: Each individual system in the ensemble is in a pure state $|\hat{\mathbf{n}}\rangle_1|-\hat{\mathbf{n}}\rangle_2$ for some specific direction $\hat{\mathbf{n}}$. The whole ensemble is simply a classical mixture of these quantum mechanical product states, with an isotropic distribution of the directions $\hat{\mathbf{n}}$. The expectation value of the product of the $\hat{\mathbf{n}}_1$ component of spin of particle 1 and the $\hat{\mathbf{n}}_2$ component of spin of particle 2 in this ensemble is obtained by first calculating

$$E_{\hat{\mathbf{n}}}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2) \equiv {}_2\langle -\hat{\mathbf{n}}|_1\langle\hat{\mathbf{n}}|(\hat{\mathbf{n}}_1\cdot\boldsymbol{\sigma}_1)(\hat{\mathbf{n}}_2\cdot\boldsymbol{\sigma}_2)|\hat{\mathbf{n}}\rangle_1|-\hat{\mathbf{n}}\rangle_2, \tag{B7}$$

and then averaging this result with an isotropic distribution over the $\hat{\mathbf{n}}$'s. Since the result will clearly only depend upon the angle between $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$, we make the simplest choice of directions making an angle α to each other:

$$\hat{\mathbf{n}}_1 = \hat{\mathbf{z}}, \quad \hat{\mathbf{n}}_2 = \cos\alpha\hat{\mathbf{z}} + \sin\alpha\hat{\mathbf{x}}. \tag{B8}$$

Using Eqs. (A2a) and (A2b) of Appendix A, we obtain

$$E^{\hat{\mathbf{n}}}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2) = -\cos^2\theta\cos\alpha - \cos\theta\sin\theta\cos\phi\sin\alpha, \tag{B9}$$

where θ and ϕ are the polar and azimuthal angles of $\hat{\mathbf{n}}$. Averaging over $\hat{\mathbf{n}}$ with an isotropic distribution yields the ensemble average

$$E^{(\text{isotropic})}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2) = -\tfrac{1}{3}\cos\alpha = -(\tfrac{1}{3})\hat{\mathbf{n}}_1\cdot\hat{\mathbf{n}}_2, \tag{B10}$$

in agreement with Eq. (11) of Bell's paper in Ref. 1. Note that even when ($\hat{\mathbf{n}}_1 = \hat{\mathbf{n}}_2$) the correlation given by Eq. (B10) is not perfect, and all types of spin outcomes ($++,+-,-+,--$) will occur.

APPENDIX C: PROOF OF BELL'S INEQUALITY

For any direction $\hat{\mathbf{n}}$ and for all λ (except for a set of probability measure zero[12]) there is the perfect correlation

$$A_\lambda(\hat{\mathbf{n}}) = -B_\lambda(\hat{\mathbf{n}}). \tag{C1}$$

Hence we can rewrite the expectation value $E^\rho(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2)$ of Eq. (4) as

$$E^\rho(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2) = -\int_\Lambda A_\lambda(\hat{\mathbf{n}}_1)A_\lambda(\hat{\mathbf{n}}_2)d\rho. \tag{C2}$$

Hence,

$$\begin{aligned} E^\rho(\hat{\mathbf{a}},\hat{\mathbf{b}}) - E^\rho(\hat{\mathbf{a}},\hat{\mathbf{c}}) &= -\int_\Lambda [A_\lambda(\hat{\mathbf{a}})A_\lambda(\hat{\mathbf{b}}) - A_\lambda(\hat{\mathbf{a}})A_\lambda(\hat{\mathbf{c}})]d\rho \\ &= \int_\Lambda [-A_\lambda(\hat{\mathbf{a}})A_\lambda(\hat{\mathbf{b}})][1 - A_\lambda(\hat{\mathbf{b}})A_\lambda(\hat{\mathbf{c}})]d\rho. \end{aligned} \tag{C3}$$

For all λ, $-A_\lambda(\hat{\mathbf{a}})A_\lambda(\hat{\mathbf{b}})$ is either $+1$ or -1 and hence has absolute value 1, and $[1 - A_\lambda(\hat{\mathbf{b}})A_\lambda(\hat{\mathbf{c}})]$ is nonnegative and therefore equals its absolute value. Therefore, taking the absolute values of the terms in Eq. (C3) yields

$$\begin{aligned} |E^\rho(\hat{\mathbf{a}},\hat{\mathbf{b}}) - E^\rho(\hat{\mathbf{a}},\hat{\mathbf{c}})| &\leqslant \int_\Lambda [1 - A_\lambda(\hat{\mathbf{b}})A_\lambda(\hat{\mathbf{c}})]d\rho \\ &= 1 + E^\rho(\hat{\mathbf{b}},\hat{\mathbf{c}}). \end{aligned} \tag{C4}$$

where use has again been made of Eq. (4) and also of the normalization condition

$$\int_\Lambda d\rho = 1. \tag{C5}$$

Hence, inequality (6) of Sec. II is proved.

APPENDIX D: BELL'S MODEL

In Bell's illustrative local model for pairs of spin-1/2 particles with total spin angular momentum zero, the space Λ of complete states consists of unit vectors in three-dimensional space, denoted by $\hat{\boldsymbol{\lambda}}$. The functions $A_{\hat{\lambda}}(\hat{\mathbf{n}})$ and $B_{\hat{\lambda}}(\hat{\mathbf{n}})$ are defined as follows:

$$\text{If } \hat{\boldsymbol{\lambda}}\cdot\hat{\mathbf{n}} \neq 0, \text{ then } A_{\hat{\lambda}}(\hat{\mathbf{n}}) = -B_{\hat{\lambda}}(\hat{\mathbf{n}}) = \text{sign}(\hat{\boldsymbol{\lambda}}\cdot\hat{\mathbf{n}}). \tag{D1}$$

(In other words, $A_{\hat{\lambda}}(\hat{\mathbf{n}}) = 1$ if $\hat{\boldsymbol{\lambda}}$ and $\hat{\mathbf{n}}$ are in the same hemisphere, and $B_{\hat{\lambda}}(\hat{\mathbf{n}}) = 1$ if $\hat{\boldsymbol{\lambda}}$ and $\hat{\mathbf{n}}$ are opposite in hemispheres.)

$$\begin{aligned} &\text{If } \hat{\boldsymbol{\lambda}}\cdot\hat{\mathbf{n}} = 0, \text{ then } A_{\hat{\lambda}}(\hat{\mathbf{n}}) \\ &\quad = -B_{\hat{\lambda}}(\hat{\mathbf{n}}) = \text{sign} \\ &\quad (\text{the first nonzero term of } n_x,n_y,n_z). \end{aligned} \tag{D2}$$

[Actually, Bell does not give a rule for the case of $\hat{\boldsymbol{\lambda}}\cdot\hat{\mathbf{n}} = 0$, but we have written (D2), as he himself does elsewhere,[16] in order to ensure that $A_{\hat{\lambda}}(\hat{\mathbf{n}})$ and $B_{\hat{\lambda}}(\hat{\mathbf{n}})$ are defined for all $\hat{\lambda}\in\Lambda$.] From Eq. (3) one sees that there are perfect correlations of the outcomes of measuring the $\hat{\mathbf{n}}_1$ spin component of particle 1 and the $\hat{\mathbf{n}}_2$ spin component of particle 2 if and only if $\hat{\mathbf{n}}_1 = \hat{\mathbf{n}}_2$ or $\hat{\mathbf{n}}_1 = -\hat{\mathbf{n}}_2$. In the first case, the outcomes are opposite; in the second, they are the same. Clearly (D1) and (D2) reproduce these perfect correlations.

APPENDIX E: DERIVATION OF THE INITIAL STATE OF THE SYSTEM IN FOUR SPIN-1/2 PARTICLES

It is assumed that a spin-1 particle decays into two spin-1 particles, I and II, each of which decays into two spin-1/2 particles, and these four are labeled 1,2,3,4. Throughout the process the total spin S is assumed to remain 1 and the total z component M is assumed to remain 0. We shall denote the spin states of I and II by the kets $|m\rangle_{\text{I}}$ and $|m\rangle_{\text{II}}$, respectively, where m is the component of spin along the specified polar axis; in both of these kets the total spin quantum number 1 has been suppressed. The spin state of the composite system consisting of I and II will be denoted by $\|S\,M\rangle\rangle$, suppressing the total spin-1 of each of the particles I and II. The spin state of the composite system after the first decay is $\|1\,0\rangle\rangle$, and it can be expressed as

$$\begin{aligned} \|10\rangle\rangle = c_1|1\rangle_{\text{I}}|-1\rangle_{\text{II}} + c_2|0\rangle_{\text{I}}|0\rangle_{\text{II}} \\ + c_3|-1\rangle_{\text{I}}|1\rangle_{\text{II}}. \end{aligned} \tag{E1}$$

The coefficients c_1,c_2,c_3 can be found in a table of Clebsch–Gordon coefficients, but we can determine them here by a shortcut. We know that all the kets $\|2M\rangle\rangle$ are symmetric under exchange of particles I and II because $\|22\rangle\rangle = |1\rangle_{\text{I}}|1\rangle_{\text{II}}$ is symmetric and the lowering operator $S_- \equiv (s_{\text{I}x} + s_{\text{II}x}) - i(s_{\text{I}y} + s_{\text{II}y})$ clearly preserves symmetry. Hence $\|11\rangle\rangle$ must be antisymmetric in order to ensure orthogonality to $\|21\rangle\rangle$, and therefore by lowering one obtains an antisymmetric $\|10\rangle\rangle$. Hence in Eq. (E1) we obtain

$$c_2 = 0, \quad c_1 = -c_3 = 1/\sqrt{2} \tag{E2}$$

(the last step by normalization). At the second decay we have

$$|1\rangle_{\text{I}} \to |1/2\rangle_1|1/2\rangle_2, \tag{E3a}$$

$$|-1\rangle_{\text{I}} \to |-1/2\rangle_1|-1/2\rangle_2, \tag{E3b}$$

$$|1\rangle_{\text{II}} \to |1/2\rangle_3|1/2\rangle_4, \tag{E3c}$$

$$|-1\rangle_{\text{II}} \to |-1/2\rangle_3|-1/2\rangle_4. \tag{E3d}$$

Hence the state of the quadruple of spin-1/2 particles 1,2,3,4 is

$$\begin{aligned} |\Psi\rangle = (1/\sqrt{2})[|+\rangle_1|+\rangle_2|-\rangle_3|-\rangle_4 \\ -|-\rangle_1|-\rangle_2|+\rangle_3|+\rangle_4], \end{aligned} \tag{E4}$$

where we have simplified the notation $|\pm 1/2\rangle$ to $|\pm\rangle$.

APPENDIX F: STATISTICAL CORRELATIONS IN THE SYSTEM OF FOUR SPIN-1/2 PARTICLES

The expectation value in the state $|\Psi\rangle$ of Eq. (7) of the product of outcomes of the $\hat{\mathbf{n}}_1$ component of spin of particle 1, the $\hat{\mathbf{n}}_2$ component of spin of particle 2, etc. is

$$\begin{aligned} E^{\Psi}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2,\hat{\mathbf{n}}_3,\hat{\mathbf{n}}_4) &= \langle\Psi|(\hat{\mathbf{n}}_1\cdot\boldsymbol{\sigma}_1)(\hat{\mathbf{n}}_2\cdot\boldsymbol{\sigma}_2)(\hat{\mathbf{n}}_3\cdot\boldsymbol{\sigma}_3)(\hat{\mathbf{n}}_4\cdot\boldsymbol{\sigma}_4)|\Psi\rangle \\ &= [\tfrac{1}{2}\langle + + - - |(\hat{\mathbf{n}}_1\cdot\boldsymbol{\sigma}_1)(\hat{\mathbf{n}}_2\cdot\boldsymbol{\sigma}_2)(\hat{\mathbf{n}}_3\cdot\boldsymbol{\sigma}_3)(\hat{\mathbf{n}}_4\cdot\boldsymbol{\sigma}_4)| + + - - \rangle \\ &\quad - \langle + + - - |(\hat{\mathbf{n}}_1\cdot\boldsymbol{\sigma}_1)(\hat{\mathbf{n}}_2\cdot\boldsymbol{\sigma}_2)(\hat{\mathbf{n}}_3\cdot\boldsymbol{\sigma}_3)(\hat{\mathbf{n}}_4\cdot\boldsymbol{\sigma}_4)| - - + + \rangle \\ &\quad - \langle - - + + |(\hat{\mathbf{n}}_1\cdot\boldsymbol{\sigma}_1)(\hat{\mathbf{n}}_2\cdot\boldsymbol{\sigma}_2)(\hat{\mathbf{n}}_3\cdot\boldsymbol{\sigma}_3)(\hat{\mathbf{n}}_4\cdot\boldsymbol{\sigma}_4)| + + - - \rangle \\ &\quad + \langle - - + + |(\hat{\mathbf{n}}_1\cdot\boldsymbol{\sigma}_1)(\hat{\mathbf{n}}_2\cdot\boldsymbol{\sigma}_2)(\hat{\mathbf{n}}_3\cdot\boldsymbol{\sigma}_3)(\hat{\mathbf{n}}_4\cdot\boldsymbol{\sigma}_4)| - - + + \rangle]. \end{aligned} \tag{F1}$$

But

$$\langle + |(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})| + \rangle = (1 \quad 0)\begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{-i\phi} & -\cos\theta \end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \cos\theta, \tag{F2a}$$

and

$$\langle + |(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})| - \rangle = (1 \quad 0)\begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix} = \sin\theta e^{-i\phi}. \tag{F2b}$$

Likewise,

$$\langle - |(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})| + \rangle = \sin\theta e^{i\phi} \tag{F2c}$$

and

$$\langle - |(\hat{\mathbf{n}}\cdot\boldsymbol{\sigma})| - \rangle = -\cos\theta. \tag{F2d}$$

The first and the last terms in the brackets of Eq. (F1) are products of four factors like those of Eqs. (F2a) and (F2d), with net result $\cos\theta_1\cos\theta_2\cos\theta_3\cos\theta_4$. The second and third terms add to $\sin\theta_1\sin\theta_2\sin\theta_3\sin\theta_4\cos(\phi_1+\phi_2-\phi_3-\phi_4)$. Collecting terms we obtain

$$E^{\Psi}(\hat{\mathbf{n}}_1,\hat{\mathbf{n}}_2,\hat{\mathbf{n}}_3,\hat{\mathbf{n}}_4) = \cos\theta_1\cos\theta_2\cos\theta_3\cos\theta_4 - \sin\theta_1\sin\theta_2\sin\theta_3\sin\theta_4\cos(\phi_1+\phi_2-\phi_3-\phi_4). \tag{F3}$$

APPENDIX G: STATISTICAL CORRELATIONS IN THE THREE-PARTICLE INTERFEROMETER

The evolution of the state $|\Psi\rangle$ of Eq. (19) can be obtained by using the evolutions (20a) and (20b) of the states of particle 1 and the analogous evolutions of states of particles 2 and 3. The result is

$$\begin{aligned} |\Psi\rangle \to \tfrac{1}{4}[&(1 - ie^{i(\phi_1+\phi_2+\phi_3)})|d\rangle_1|e\rangle_2|f\rangle_3 + (i - e^{i(\phi_1+\phi_2+\phi_3)})|d\rangle_1|e\rangle_2|f'\rangle_3 \\ &+ (i - e^{i(\phi_1+\phi_2+\phi_3)})|d\rangle_1|e'\rangle_2|f\rangle_3 + (-1 + ie^{i(\phi_1+\phi_2+\phi_3)})|d\rangle_1|e'\rangle_2|f'\rangle_3 \\ &+ (i - e^{i(\phi_1+\phi_2+\phi_3)})|d'\rangle_1|e\rangle_2|f\rangle_3 + (-1 + ie^{i(\phi_1+\phi_2+\phi_3)})|d'\rangle_1|e\rangle_2|f'\rangle_3 \\ &+ (-1 + ie_i^{\,i(\phi_1+\phi_2+\phi_3)})|d'\rangle_1|e'\rangle_2|f\rangle_3 + (-i + e^{i(\phi_1+\phi_2+\phi_3)})|d'\rangle_1|e'\rangle_2|f'\rangle_3]. \end{aligned} \tag{G1}$$

The probability for detection of the three particles by the respective detectors d,e,f is

$$\begin{aligned} P^{\Psi}_{def}(\phi_1,\phi_2,\phi_3) &= \tfrac{1}{16}|(1 - ie^{i(\phi_1+\phi_2+\phi_3)})|^2 \\ &= \tfrac{1}{8}[1 + \sin(\phi_1+\phi_2+\phi_3).] \end{aligned} \tag{G2a}$$

Likewise,

$$P^{\Psi}_{d'ef}(\phi_1,\phi_2,\phi_3) = \tfrac{1}{8}[1 - \sin(\phi_1+\phi_2+\phi_3)]. \tag{G2b}$$

The difference between Eqs. (G2a) and (G2b) is exactly what one would expect, since replacing d by d' has the effect of requiring the beam a to be reflected (rather than transmitted) from the beam splitter in order to reach the designated detector and requiring beam a' to be transmitted (rather than reflected) in order to reach it. The replacement thus removes an i from one of the two terms in the coefficient of the ket $|d\rangle_1|e\rangle_2|f\rangle_3$ and inserts an i in the other term. Hence, there is a net change of sign in the cross terms when the absolute square is taken. Iterating this argument shows that the probability that a triple of particles will cause three designated detectors to fire will be the same as for d,e,f [given by Eq. G2a] if an even number of the detectors are primed; and the probability will be the same as for d',e,f [given by Eq. (G2b)] if an odd number of them are primed.

If detection by an unprimed detector (for example, d) is assigned the value $+1$ and by a primed detector (for example, d') is assigned the value -1, then the expectation value of the product of the outcomes is

$$\begin{aligned} E^{\Psi}(\phi_1+\phi_2+\phi_3) = P^{\Psi}_{def}(\phi_1,\phi_2,\phi_3) &+ P^{\Psi}_{def'}(\phi_1,\phi_2,\phi_3) + P^{\Psi}_{d'ef'}(\phi_1,\phi_2,\phi_3) + P^{\Psi}_{d'e'f}(\phi_1,\phi_2,\phi_3) - P^{\Psi}_{def'}(\phi_1,\phi_2,\phi_3) \\ &- P^{\Psi}_{de'f}(\phi_1,\phi_2,\phi_3) - P^{\Psi}_{d'ef}(\phi_1,\phi_2,\phi_3) - P^{\Psi}_{d'e'f'}(\phi_1,\phi_2,\phi_3). \\ = \sin(\phi_1+\phi_2+\phi_3). \end{aligned} \tag{G3}$$

[a] Dedicated to the memory of John S. Bell.

[1] J. S. Bell, "On the Einstein–Podolsky–Rosen paradox," Physics 1, 195–200 (1964), reprinted in J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge U.P., Cambridge, 1987).

[2] A. Einstein, B. Podolsky, and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?" Phys. Rev. **47**, 777–780 (1935).

[3] The earliest was by S. J. Freedman and J. S. Clauser, "Experimental test of local hidden-variable theories," Phys. Rev. Lett. **28**, 938–941 (1972); A. Aspect, J. Dalibard, and G. Roger, "Experimental tests of Bell's inequalities using time-varying analyzers," Phys. Rev. Lett. **47**, 1804–1807 (1982) reports an experimental test, in which rapid switches were made among polarization analyzers in an effort to prevent subluminal communication. A summary of experiments up to 1987 is given by M. Redhead, *Incompleteness, Nonlocality, and Realism* (Clarendon, Oxford, 1987), pp. 107–113.

[4] D. M. Greenberger, M. Horne, and A. Zeilinger, "Going beyond Bell's theorem," in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, edited by M. Kafatos (Kluwer Academic, Dordrecht, The Netherlands, 1989), pp. 73–76.

[5] This proof was never published by S. Kochen, but was reported by him orally to one of the authors (A.S.) in the early 1970's. P. Heywood and M. I. G. Redhead, "Nonlocality and the Kochen–Specker paradox," Found. Phys. **13**, 481–499 (1983), mention Kochen's proof in a footnote but seem to have discovered it independently. The proof will be summarized below in Ref. 16.

[6] D. Bohm, *Quantum Theory* (Prentice-Hall, Englewood Cliffs, NJ, 1951), pp. 614–623.

[7] D. M. Greenberger and H. S. Choi, presentation at the Nineteenth Workshop on Quantum Optics, Snowbird, Utah, January 1989 (unpublished).

[8] N. D. Mermin, "Quantum mysteries revisited," Am. J. Phys. **58**, 731–734 (1990); and "What's wrong with these elements of reality?" Phys. Today **43**(6), 9–11 (1990).

[9] M. A. Horne and A. Zeilinger, "A Bell-type EPR experiment using linear momenta," in *Symposium on the Foundations of Modern Physics*, edited by P. Lahti and P. Mittelstaedt (World Scientific, Singapore, 1985), pp. 435–439; R. Ghosh and L. Mandel, "Observation of nonclassical effects in the interference of two photons," Phys. Rev. Lett. **59**, 1903–5 (1987); M. A. Horne, A. Shimony, and A. Zeilinger, "Two-particle interferometry," Phys. Rev. Lett. **62**, 2209–2212 (1989); and many other papers by L. Mandel and collaborators.

[10] There is a subtle point in EPR's argument that deserves comment. In the reality assumption (iii) the phrase "can predict" occurs. The phrase is ambiguous, because it may be understood in the strong sense, that data are at hand for making the prediction, or in the weak sense, that a measurement could be made to provide data for the prediction. EPR assume the weak sense, and indeed unless they did so they could not argue that an element of physical reality exists for all component of spin, those *which could have been measured* as well as *the one that actually was measured*. Bohr's famous answer to EPR ["Can quantum-mechanical description of physical reality be considered complete?" Phys. Rev. **48**, 696–702 (1935)] can be read as a challenge to EPR on this point, maintaining that the reality assumption (iii) is correct only if the strong sense of "can predict" is used. The preference for one rather than the other of these two interpretations of the phrase is not merely a semantical matter, but is an indication of a philosophical commitment. Bohr believed that the concept of "reality" cannot be applied legitimately to a property unless there is an experimental arrangement for observing it, whereas Einstein regarded this view as anthropocentric and maintained that physical systems have intrinsic properties whether they are observed or not. For an analysis of their views, see A. Shimony, "Physical and philosophical aspects of the Bohr–Einstein debate," *Neils Bohr: Physics and the World*, edited by H. Feshbach *et al.* (Harwood, Chur, 1988).

[11] Contrary to the opinion of numerous expositors, Bell did not begin with an assumption of a deterministic hidden variables theory. He follows EPR, and since their argument concludes that there are elements of physical reality that predetermine the outcomes of measurements of spin of particles 1 and 2, "determinism" need not be assumed as an extra premiss; and since these elements of physical reality are not contained in the quantum state, they are "hidden variables," if that phrase refers to any physical reality missing from the quantum description. These points are stressed by R. Clifton, M. Redhead, and J. Butterfield, "Generalization of the Greenberger–Horne–Zeilinger algebraic proof of nonlocality" (to be published).

[12] Two technical points concerning probability measure need to be mentioned. First, the set $\Lambda_{\hat{n}}$ was introduced above informally as the set of λ 's agreeing with the perfect correlation of Eq. (3a) for a specific $\hat{\mathbf{n}}$. Since the correlation is stated as an expectation value, $\Lambda_{\hat{n}}$ is determined only within a set of probability measure zero. However, nothing is lost, and there is a gain in simplicity, if we merely identify $\Lambda_{\hat{n}}$ with the set of all complete states that predetermine that spin measurements (for specific $\hat{\mathbf{n}}$) on the two particles have opposite outcomes. Second, the argument for the existence of the outcome functions $A_\lambda(\hat{\mathbf{n}})$ and $B_\lambda(\hat{\mathbf{n}})$ does not prove that these functions are well defined if λ does not belong to the set $\Lambda_{\hat{n}}$, but the integrand of the right-hand side of Eq. (4) is written as if it is well defined for any λ in the space of complete states Λ. This difficulty is easily resolved by assigning arbitrarily a value $+1$ or -1 to $A_\lambda(\hat{\mathbf{n}})$ and to $B_\lambda(\hat{\mathbf{n}})$ when λ belongs to $\overline{\Lambda}_{\hat{n}}$ (the complement of $\Lambda_{\hat{n}}$). Since $\rho(\overline{\Lambda}_{\hat{n}})$ is zero, the values of the integrand for $\lambda \in \overline{\Lambda}_{\hat{n}}$ make no difference to the integral.

[13] Although not indicated in Fig. 2, the angle between the diverging beams 2 and 2 (3 and 4) is assumed small, so that the *orbital* angular momentum need not be considered during the *spin* calculation of Eq. (E7) in Appendix E. Also, Fig. 2 does not exhibit the two intermediate particles I and II, whose decay, as discussed in Appendix E, produces the four final particles 1,2,3,4; the production and decay of particles I and II may be imagined to occur inside the "source" sphere of Fig. 2.

[14] Agreement with the perfect quantum mechanical correlations (10a) or (10b) for a specific $\phi_1,\phi_2,\phi_3,\phi_4$ requires that (11a) or (11b) hold only for a set of λ 's of probability measure unity (cf. Ref. 12). Hence, to be mathematically rigorous, we should say that (11a) and (11b) and all the equations through Eq. (18) hold "almost everywhere" (i.e., except for a set of λ 's of probability measure zero). It is important to observe that if each of a finite (or even countable) set of statements holds almost everywhere, their conjunction does also, because the finite (or countable) union of sets of measure zero also has measure zero. It is not always the case that the uncountable union of sets of measure zero is a set of measure zero, as Clifton, Redhead, and Butterfield warn in the paper cited in Ref. 11; but the argument from (11a) and (11b) to (18) does not require taking the union of an uncountable set of sets.

[15] N. D. Mermin (private communication) has noted a shortened derivation of Eq. (16) from Eqs. (12). Specifically, just multiply the first three equations of (12) and compare the result with the fourth to conclude Eq. (16). We maintain the proof given in the text because it was the original proof presented at conferences and has been commented upon.

[16] It is noteworthy that the arguments of Kochen and Heywood and Redhead, mentioned in Ref. 5, also depend upon using a system more complex than a pair of spin-1/2 particles. They consider a pair of spin-1 particles with total spin angular momentum zero. Applying EPR's premisses (ii), (iii), and (iv) leads to the conclusion that the value ($+1$, 0, or -1) of any spin component of either particle is an element of physical reality, and hence to the weaker conclusion that the value ($+1$ or 0) of the *square* of any spin component is an element of physical reality. We focus on this weaker conclusion because if $\hat{\mathbf{n}}_1$, $\hat{\mathbf{n}}_2$, $\hat{\mathbf{n}}_3$ are orthogonal directions, then $s^2_{\hat{n}_1}$, $s^2_{\hat{n}_2}$, $s^2_{\hat{n}_3}$ are commuting operators (even though $s_{\hat{n}_1}$, $s_{\hat{n}_2}$, $s_{\hat{n}_3}$ are not) and hence in principle are simultaneously measurable according to quantum mechanics. In addition, quantum mechanics predicts for the spin-1 particle that the sum $s^2_{\hat{n}_1} + s^2_{\hat{n}_2} + s^2_{\hat{n}_3}$ has the definite value 2 in any state. Hence, the conjunction of the quantum mechanical predictions with EPR's premisses (ii), (iii), and (iv) leads to the following assertion concerning the spin-1 particle: There exist definite values 1 or 0 for every squared spin component $s^2_{\hat{n}}$, such that the values assigned to $s^2_{\hat{n}_1}$, $s^2_{\hat{n}_2}$, $s^2_{\hat{n}_3}$ sum to 2 if $\hat{\mathbf{n}}_1$, $\hat{\mathbf{n}}_2$, $\hat{\mathbf{n}}_3$ are mutually orthogonal. But this assertion is impossible, as shown in different ways (but all quite elaborate) by A. Gleason, "Measures on the closed subspaces of a Hilbert space," J. Math. Mech. **6**, 885–893 (1957); by J. S. Bell, "On the problem of hidden variables in quantum mechanics," Rev. Mod. Phys. **38**, 447–452 (1966), and reprinted in his book mentioned in Ref. 1; by S. Kochen and E. Specker, "The problems of hidden variables in quantum mechanics," J. Math. Mech. **17**, 59–88

(1967); and E. de Obaldia, A. Shimony, and F. Wittel, "Amplification of Belinfante's argument for the nonexistence of dispersion-free states," Found. Phys. **18**, 1013–1021 (1988).

[17] See the paper of Horne and Zeilinger of Ref. 9; also M. Horne, A. Shimony, and A. Zeilinger "Introduction to two-particle interferometry," in *Sixty-Two Years of Uncertainty*, edited by A. Miller (Plenum, New York, 1990); and J. G. Rarity and P. R. Tapster, "Experimental violation of Bell's inequality based on phase and momentum," Phys. Rev. Lett. **64**, 2495–2498 (1990) (the first actual performance of a spinless test of Bell's inequality).

[18] If the particles 1,2,3 are photons, then the 50–50 beam splitter can be a half-silvered mirror, and a glass plate of appropriate thickness can serve as a phase plate.

[19] The phase relation given in (20a) and (20b) is actually only a special case of the phase relations allowed by unitarity between the transmitted and reflected beams emerging from an ideal symmetric beam splitter, but all allowable phase relations lead to essentially the same predictions [the only difference being the addition of a fixed phase angle to $\phi_1 + \phi_2 + \phi_3$ in Eqs. (21a) and (21b)]. This statement can be verified by paralleling an argument in the paper of Horne, Shimony, and Zeiliner of Ref. 17. See also A. Zeilinger, "General properties of lossless beam splitters in interferometry," Am. J. Phys. **49**, 882–883 (1981).

[20] Clearly an analogous generalization can be carried out for the spin correlations considered in Sec. III.

[21] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed experiment to test local hidden-variable theories," Phys. Rev. Lett. **23**, 880–884 (1969).

[22] N. D. Mermin, "The EPR experiment—thoughts about the 'Loophole,'" *New Techniques and Ideas in Quantum Measurement Theory*, edited by D. M. Greenberger (New York Academy of Science, New York, 1986), pp. 422–428.

[23] In reaching the conclusion that η must exceed 0.908 in order to rule out the EPR program, we have avoided *all* auxiliary assumptions. N. D. Mermin arrived at the same limit by an independent argument. This limit can be significantly reduced by plausible assumptions, such as symmetry properties of the result functions *A*, *B*, *C* (to be published).

[24] J. S. Bell, "Introduction to the hidden-variable question," *Foundations of Quantum Mechanics*, edited by B. d'Espagnat (Academic, New York, 1971), pp. 171–181. Reprinted in his book mentioned in Ref. 1.

[25] J. F. Clauser and M. A. Horne, "Experimental consequences of objective local theories," Phys. Rev. D **10**, 526–535 (1974).

[26] H. P. Stapp, "Locality and reality," Found. Phys. **10**, 767–795 (1980).

[27] A. Fine, "Joint distributions, quantum correlations, and commuting observables," J. Math. Phys. **23**, 1306–1310 (1982).

[28] C. A. Kocher and E. D. Commins, "Polarization correlation of photons emitted in an atomic cascade," Phys. Rev. Lett. **18**, 575–577 (1967).

[29] D. C. Burnham and D. L. Weinberg, "Observation of simultaneity in parametric production of optical photon pairs," Phys. Rev. Lett. **25**, 84–87 (1970).

[30] The most general three-photon *polarization* state is an arbitrary superposition of the eight product states $|R\rangle_1|R\rangle_2|R\rangle_3$, $|R\rangle_1|R\rangle_2|L\rangle_3$, $|R\rangle_1|L\rangle_2|R\rangle_3$, etc., where $|R\rangle_i$ and $|L\rangle_i$ denote right and left circular polarization states of the *i*th photon. Since we have assumed that *spatially* the three photons propagate with 120° separation within a plane, only the two terms appearing in Eq. (33) have the required total angular momentum zero.

[31] J. D. Franson, "Bell inequality for position and time," Phys. Rev. Lett. **62**, 2205–2208 (1989).

[32] Z. Y. Ou, X. Y. Zou, L. J. Wang, and L. Mandel, "Observation of nonlocal interference in separated photon channels" Phys. Rev. Lett. **65**, 321–324 (1990).

[33] P. G. Kwiat, W. A. Vereka, C. K. Hong, H. Nathel, and R. Y. Chiao, "Observation of nonlocal two-photon interference in a dual-beam Michelson interferometer," Phys. Rev. A **41**, 2910–2913 (1990).

Electron wavelike behavior: A historical and experimental introduction

G. Matteucci
Centro di Microscopia Elettronica and GNSM-CNR, Dip. di Fisica, Università di Bologna, Via Irnerio 46, 40126 Bologna, Italy

(Received 13 November 1989; accepted for publication 7 March 1990)

Following the Fresnel theory of light and the consequent scientific debate that his formulation generated, two key experiments with electrons, diffraction by means of a circular hole and a circular obstruction, have been realized to show the existence of the Fresnel zones and of the so-called "Poisson spot." The basic arguments concerning the quantum mechanical nature of electrons can be introduced by taking advantage of the vivid impression stimulated by the experimental images.

I. INTRODUCTION

The debate that led the scientific community to accept the wave behavior of what we call light constitutes one of the most fascinating parts of optics.

The conclusive facts are well known.

In an extremely fruitful paper presented at the end of July 1818 for participation in a public competition organized by the French Academy, Fresnel, the major "architect" of the wave theory of light, described diffraction by a border and interference phenomenon occurring in the shadow of a thin wire. He gave a mathematical and physical interpretation by combining the principle of elementary wavelets with that of interference.

The Newtonian concept of light, nevertheless, was deep rooted, thus the Fresnel paper became a new subject for heated controversies. A member of the judging committee, Poisson, considering the integrals reported in the Fresnel paper, deduced the singular and "unbelievable" result that even behind a circular obstacle there should be light when this latter is illuminated by a beam of light rays. The chairman of the committee, Arago, performed this experiment and the observation confirmed the calculations. This unexpected result that, by *reductio ad absurdum*, should have

