Interference of light and Bell's theorem

A. V. Belinskiĭ and D. N. Klyshko

M. V. Lomonosov Moscow State University
(Submitted 13 March 1993)
Usp. Fiz. Nauk **163**, 1–45 (August 1993)

A review is presented of light-intensity interference experiments demonstrating the inadequacy of the classical concept of *a priori* properties of quantum objects (Bell's theorem). A detailed description of experimental procedures using N-photon light ($N=2,3,.....$) is given and is accompanied by interpretations of experimental data in terms of quantum theory and a variety of clear classical models. Different formulations of the conflict between classical ideas and quantum theory are presented, including the Bell theorem without inequalities, the Kochen–Specker theorem, and so on.

Dedicted to the memory of
Sergeĭ Aleksandrovich Akhmanov

1. INTRODUCTION

The problems formulated many years ago by Einstein, Podolsky, and Rosen (EPR),[1] Bohm,[2] and Bell[3] (see also the review literature[4–10] have continued to disturb new generations of physicists. This has largely been due to the fact that the contradiction found by Bell between the predictions of quantum theory (QT) and the hidden variable theory (HVT) can be more or less satisfactorily removed (in favor of QT, of course) by a critical experiment, which distinguishes it from other quantum paradoxes. HVT is closely related to the statistical ensemble interpretation of QT, so that such experiments—actual or thought—constitute a serious contribution to the on-going dispute between the supporters of statistical (Einstein) and orthodox (Bohr or Copenhagen) interpretations and their numerous modifications (see the review paper in Ref. 8 and Sudburys textbook.[10]

It is possible that this dispute will be resolved in the future (possibly in favor of some new interpretation) and that the historians of physics will look upon it as another striking example of a delusion that had trapped even the ablest minds in the past.

Several new lines of theoretical and experimental research have emerged in this field in recent years. Our aim in this review is to describe those that, in our view, are of particular interest. They include the following:

—the use of efficient nonlinear-optics parametric light sources that produce directed beams of 'biphotons' (correlated pairs of practically simultaneously created photons) and constitute an essential element of new modifications of optical EPR-type experiments.[11–13]

—the development of three- and generally, N-channel model experiments with EPR-Bohm type correlation,[14,15] and formulation of the corresponding generalized Bell inequalities (BI) in the form given by Mermin *et al.*[16–21]

—the Greenberger–Horne–Zeilinger theorem (GHZ) or the Bell theorem without inequalities[14,15,22,23]

—simple new examples[24,25] demonstrating and proving the Kochen–Specker theorem (KS)[26–29]

The implementation of optical methods of varifying BI, first achieved by Clauser *et al.*[30–32] (see also the reviews in Refs. 4, 5, and 9) has a number of practical advantages and can be simply described, so that clear classical and semiclassical models can be developed. Indeed, the possibility of parallel quantum-mechanical and classical treatments of optical experiments forms the basis of our presentation. We hope that this approach will facilitate at least partial reconciliaton of the quantum-mechanical and classical interpretations of the essence of EPR-Bohm type experiments and, generally, multiphoton interference effects. The same purpose is served by our preference for the Heisenberg representation which we use in the quantum-mechanical description of interference. We use it in order to demonstrate the infelicity of the word nonlocality which is almost always looked upon as a decisive symptom of these quantum effects.

The Heisenberg representation enables us to transfer quantum paradoxes (at any rate, optical paradoxes) from the output of the interferometer to its input. Together with its detectors, the interferometer can be regarded as a classical device for the investigation of the true quantum mechanical object, i.e., the original field. In this way, only the statistical properties of the still-localized field at the input are nonclassical, and the problem reduces to the basic QT paradox, namely, the principle of complementarity. The ideas that we will present will apply to the currently popular two- and many-photon interference effects; they are not directly related to BI, but are widely regarded as demonstrating nonlocality.[33–35]

Although the EPR and Bell paradoxes are described in the popular literature (see, for example, Refs. 36–38) and even in textbooks on quantum mechanics (e.g., in Sudburys book,[10] which is remarkable in many respects), they nevertheless carry an aura of mystery and are regarded as

 1063-7869/93/080653-41$10.00 © 1993 American Institute of Physics

esoteric, with a reputation for inaccessibility to the uninitiated. This is partly due to the traditional use of Bohms spin model[2] and partly to the description of optical models in terms of the Schrödinger representation, i.e., in terms of the evolution of the state vector of the two-photon field as it passes through the interferometer. All this inhibits comparisons with the familiar picture of interference of classical waves.

Appendices I and III show that there is a one-to-one correspondence between N correlated half spins and optical N-photon fields. Each photon thus belongs to two modes that differ by the type of polarization and/or the direction of the wave vector. The Stern–Gerlach magnetic spin analyzer has as its analog the polarization analyzer (prism) and the beamsplitter with a phase compensator that mix the modes in pairs. The analogy can also be extended to particles with arbitrary spin j, in which case each photon belongs to $2j+1$ modes.[39]

We shall devote considerable attention to experimental procedures that are used to introduce the necessary concepts and mathematical symbols into the theory. In general, it seems that most quantum-mechanical paradoxes, and concepts such as the KS theorem and the reduction of the wave-function, are best treated operationally, i.e., in terms of particular experiments (at any rate, thought experiments) and measuring procedures, including the operation of averaging in the course of repeated sampling of identically prepared quantum-mechanical systems. This is why Secs. 3–5 begin with the description of experiments. This approach, has lead, for example, to the conclusion that, contrary to expectations, single samplings do not reveal contradictions between QT and HVT even in the case of ideal (complete) corelations between observer readings.

Our plan will be clear from the respective section headings. After a brief review of the literature, published mostly during the last decade, we systematically examine two-, three-, and N-channel EPR experiments, which are described in terms of HVT, QT and several classical models in which the elusive hidden parameters are brought into the real world as the random phases of quasimonochromatic waves generated by parametric converters.

A separate section is devoted to the KS theorem, the Stapp paradox, and the contradiction that ensues from the Cauchy–Schwartz inequality.

Additional and auxiliary material is presented in Appendices I–IV.

We have tried to make the individual sections as independent as possible. Readers wanting only a cursory glimpse of the Bell inequality and the corresponding optical experiments with two-photon light may confine their attention to Secs. 3.1–3.4 which provide a detailed description of the simplest experiment and its interpretation in terms of three different approaches employing, respectively, Bells phenomenological theory of hidden parameters, the classical model of interfering waves with random phases (which we proposed in Ref. 19), and, finally, the formal quantum-mechanical description of photon correlations. Subsequent sections and appendices are designed for those interested in a detailed examination of the problem and in modern research in this field.

2. REVIEW OF LITERATURE

In 1935, Einstein, Podolsky, and Rosen[1] considered a quantum-mechanical system consisting of two correlated particles and concluded that formal quantum theory (QT) does not provide a complete description of physical reality. Hence it follows that it is possible to introduce certain additional parameters λ that offer a complete description without any randomness. To illustrate the EPR derivaton, Bohm[2] examined a system of two half spins. In 1964, Bell[3] gave a very general proof that HVT and QT lead to mutually contradictory predictions in Bohms model. One of the few assumptions introduced by Bell was the natural assumption of *locality*, i.e., the absence of mutual influence between two distant measuring devices.

The contradiction found by Bell (this is the so-called Bell theorem or paradox) throws doubt on the EPR program (we assume that the words *contradiction, theorem, and paradox* can be regarded as synonymous in this context). In the simplest case, the contradiction arises not only in Bohms model, but also in experiments with two correlated photons, each of which belongs to two modes[1] (11–13, 30–32, 39–41). A certain combination of measured variables, S, which we shall call the *Bell variable*, cannot exceed unity after averaging over the probability distribution function ρ_λ where λ is the set of hidden parameters, i.e.,

$$|\langle S\rangle_\rho|\leqslant 1, \tag{2.1a}$$

where

$$\rho_\lambda\geqslant 0, \quad \int_{-\infty}^{\infty}\rho_\lambda \mathrm{d}\lambda=1.$$

It will be convenient to use a definition of S that is smaller by a factor of two as compared with the generally accepted definition.

On the other hand, QT contains factorizable states $|\psi\rangle$ which are called mixed or entangled states in which the expectation value of the operator S corresponding to the Bell variable takes the value

$$\langle S\rangle_\psi=\sqrt{2} \tag{2.1b}$$

(as a rule, we shall not explicitly distinguish between the observable and the corresponding operator).

The contradiction indicated by (2.1a, b), which will be derived in detail in Sec. 3, has frequently been varified experimentally, mostly by optical methods.[11–13,30–32] Of course, this involved the use of the frequency 'definition' of the expectation value (the only one possible in practice):

$$\langle S\rangle_{\exp}\equiv\frac{1}{L}\sum_{i=1}^{L}S_i, \tag{2.2}$$

where L is the number of samplings.

These experiments were essentially based on a form of the Brown–Twiss *intensity interference*[42,43] whose particu-

lar feature is the harmonic modulation of the correlation between the intensities of two light beams (modes):

$$\langle I_1 I_2 \rangle \propto 1 + V \cos \varphi, \tag{2.3}$$

where V is the visibility and φ is a particular combination of phase delays in the optical channel, or in the polarization experiment, double the angle between the analyzer axes.

The phenomenological Bell theory operates with discrete two-valued (dichotomic) observables such as $A = \pm 1$, so that the detection process is based on a photon-counting regime. QT predicts perfect visibility $V = 1$ when one photon simultaneously enters each of the interferometer inputs. In practice, this type of stationary two-photon light is generated either in two-photon transitions (casade[30,31] or direct[32]) in atomic beams or, more efficiently, by splitting the primary pump radiation into pairs of secondary photons by *parametric scattering* or parametric down-conversion of frequency.[11–13]

If photon-pair creation occurs too frequently, the pairs will occasionally overlap within the detection time constant, which will lead to random coincidences between photon counts. This will reduce the interference visibility, so that the quantum and experimental values of the Bell observable will be such that

$$\langle S \rangle = \sqrt{2} V, \tag{2.4}$$

and the BI will not be violated for $V \leqslant 0.71$.

Classical models describing interference between random waves (Sec. 3.3) are subject to more stringent visibility limit $V \leqslant 0.5$.

However, in most experiments, the visibility V exceeds the critical value 0.71. It is important to note at this juncture that we shall adopt the following convention: the validity of the Bell analysis is limited in practice by additional (and not always explicit) assumptions that are reasonable or even obvious in character, but whose physical significance must nevertheless be critically examined.[44,45]

Most EPR-Bell type experiments are based on polarization interferometry with polarization-correlated photons. However, the advent of efficient parametric sources of parallel beams of polarized photons with correlated phases, i.e., with correlated quadrature components, has lead to the emergence of new types of intensity interferometer. For example, Rarity and Tapster[13] have reached $V = 0.8$ in a system consisting of two Mach–Zender interferometers without beam-splitters at entry. The principle of this experiment was put forward a few years ago[46,47] and will be examined in detail in Sec. 3.1. Similiar modifications are discussed in Refs. 39 and 48–53.

An experiment demonstrating the EPR paradox for continuous observables in a homodyne interferometer has recently been carried out for the first time.[54] Experiments of this kind were previously discussed in Refs. 55–59 (cf. Sec. 3.7).

Other interesting applications have also recently appeared.[60–72] For example, Zukovskiĭ and Zeilinger[70] have analyzed a combined system in which one of the channels contains a polarization interferometer and the other a Mach–Zender interferometer.

Joshi and Lawande[72] have discussed a system of strongly excited two-level atoms and the resonance fluorescence emitted by them, which was used as a source of light for an intensity interferometer generating pairs of correlated photons.

Oliver and Stroud[71] suggested that, instead of spins and photons, it is possible to use two or three Rydberg atoms excited to a nonfactorizable state by one common photon. The BI introduced by them for three observers was found to be a special case of the N-channel BI found by Mermin[16] and Hardy.[20]

It was shown in Ref. 64 that parametric scattering could be used to demonstrate the EPR paradox for observables $q_\perp$ and $p_\perp$, i.e., the transverse position coordinate and momentum of two photons with correlated directions of propagation. Similar possibilities were examined in Refs. 65–67 for pairs of photon observables such as energy and time since creation.

We now turn to theoretical publications on the generalization and interpretation of BI. Bells pioneering paper[3] was followed by alternative derivations and modifications of his inequalities, which were based on the existence of a joint distribution of all the observables without explicitly mentioning the hidden parameters (see, for example, Ref. 73). For the sake of brevity, we shall not distinguish between these variants and will follow the logic and the original notation employed by Bell.

Considerable effort has been directed toward a logical analysis of the role of hidden parameters, joint distributions, locality, determinism, and so on (see Refs. 74–86). Estimates of the maximum possible violations of the BI within the framework of QT, and searches for the corresponding states, have been carried out by Cirelson[87] and by a number of other workers.[88–92]

Apart from Bell[3] (see also Ref. 15) there was also a paper by Barut and Meystre[93] who put forward clear classical models for two particles with opposite rotational angular momenta (see Appendix II).

Generalizations of BI to the case of two particles with arbitrary spin j and $N > 2$ have become popular.[102] Gisin and a number of other workers have shown that any nonfactorizable state with $N \geqslant 2$ gives rise to a violation of the BI and that the contradiction extends to arbitrarily large j (Refs. 103–106).

The Bell inequalities and quantum correlations are discussed in Refs. 107–110 from the standpoint of information theory.

A re-examination of the Kochen–Specker (KS) theorem[24,25] and this has lead to a $+1 = -1$ contradiction within the QT formalism when operators (in a set of identities consisting of commuting operators) are replaced with their eigenvalues (see Sec. 6.1). The connection between this type of paradox and Lorentz invariance is discussed in Refs. 111–114. Dewdney[115] has given a logical analysis of the basic assumption underlying the KS theorem in the Peres[24] and Mermin[25] formulations in the context of the

well known HVT approach proposed by de Broglie and Bohm. The KS contradiction is examined in Ref. 116 in the limit as $N \to \infty$, and the connection between this paradox and nonlocality is analyzed by Heywood and Redhead in Ref. 117.

Greenberger, Horne and Zeilinger (GHZ)[14,15] have considered a modification of the EPR-Bohm experiment for $N=3$ and 4 and have identified a new type of contradiction between QT and HVT (see Sec. 4.3), which also reduces to the form $+1=-1$. The GHZ paradox differs from (2.1a, b) in two principal ways: it involves neither satistical averaging nor inequalities; it is therefore also refered to as the Bell theorem without inequalities. The work of Greenberger *et al.*[14,15] immediately evoked a substantial response in the literature.[16–33,118–127] Mermin[25,118] clearly demonstrated the essence of the paradox, virtually without any formulas,[118] and has analyzed its relation to the KS theorem.[25]

The nonfactorizable states considered by GHZ were the starting point for the formulation of new and interesting thought experiments[125,126] that have given rise to considerable discussion and have stimulated further generalizations[16–20] of BI to the case of an arbitrary number N of spin 1/2 particles. The most interesting variant seems to be that found by Mermin[16] and developed further by Roy and Singh[17] and also by the present authors[19] who proposed an optical experiment (see also Ref. 21). The contradiction (2.1a, b) then takes the form of the inequality

$$|\langle S_N\rangle_\psi/\langle S_N\rangle_\rho| \geqslant 2^{(N-1)/2}. \tag{2.5}$$

In particular, $N=2$ gives (2.1a, b) whereas $N=3$ (but not $N>3$) gives Hardy's result.[20] According t Ref. 16, when $N \gg 1$, we can speak of a new macroscopic quantum effect. A new derivation of (2.5) in a sufficiently general form is given in Sec. 5 (without using the spin model).[19]

The questions that we have been discussing have recently undergone an unexpected development in cryptography, i.e., the science of coding of messages, designed to prevent eavesdropping by others.[128–131] A theoretical and experimental investigation has been made of applications of correlated photon pairs to noise-free communication, measurements of distance,[132–135] and photometry.[136–138]

We shall briefly consider the interpretation of the Bell and KS theorems that show the incompatibility between classical ideas and the quantum formalism. These theorems and the corresponding experiments that confirm QT predictions are widely thought to consititute evidence for the *nonlocality* of not only specific physical processes, but also of quantum mechanics itself. The term *nonseperability* is also occasionally used (see below for more details).

It is considered that 'EPR type nonlocality' is different from 'signal nonlocality' that implies the existence of signals that propagate with superluminal velocity by means of quantum correlations. Refutations of the latter possibility are discussed in Refs. 139–143. The connection between this question and Weinberg's nonlinear quantum mechanics is examined in Refs. 144–145; we also note that doubts have been expressed about the validity of the particle-wave duality and the Feynman's uncertainty principle.[146] However, let us now return to our main theme.

Very few publications dispute[21,147–151] the necessity of nonlocality in the resolution of the Bell, KS, and GHZ paradoxes. New arguments relating to this will be presented later.

There are two main ways of achieving formal reconciliation between HVT, on the one hand, and experiments and QT, on the other. They are: the acknowledgment of nonlocality (due to unknown interactions between quantum particle detectors) and/or the assumption of negative probability distributions for measured variables.[9,93,152–153] Either hypothesis constitutes a very heavy price to pay for the survival of 'objective realism' as an alternative to the principle of complementarity, and hardly anyone accepts them literally. The true state of affairs is obscured by the abuse of the term 'nonlocality', and we shall now try to apply some logic to the situation, which we shall formulate as follows in the case of BI.

There are two basic premises:

(a) classical local theories lead to a particular inequality

(b) quantum theory violates this inequality; in violation of the rules of formal logic, this is said to lead to the conclusion that

(c) quantum theory is nonlocal

For comparison, we now present a comparable syllogism:

(a′) all good stories are short

(b′) this novel is long conclusion:

(c′) it is a bad novel.

Sometimes, the phrase 'quantum theory' in (b) and (c) is replaced with 'observable quantum effects', but this does not, of course, alter the essense of the situation.

A clear independent definition of the precise meaning of (c) does not appear to exist, so that the only possibility that remains is that (c) is simply a more concise symbolic representation of the original premises (a) and (b).

The fact that 'nonlocality' is unfounded in QT also follows directly from the quantum-mechanical description of EPR-Bell type optical experiments in the Heisenberg representation. The latter is very close to classical statistical electrodynamics which can hardly be accused of nonlocality. In this sense, the concept of 'nonseparability', meaning quantum correlations between two or more particles (not necessarily widely separated from each other), seems somewhat more appropriate, but in mathematics its meaning is quite different.

Phrases such as 'algebraic proof of the nonlocality of QT' or 'proof of the incompatibility between QT and noncontextual HVT' are employed in connection with the KS theorem. *Contextuality* is then understood to mean that the result of a measurement of an observable A depends on which other observables B, C,... are recorded at the same time, i.e., it is again assumed that there is some hidden interaction (telepathy?) between observers or instruments recording the variables A, B, C,... .

Such interpretations have no independent meaning or

definition outside the KS theorem itself. The essense of this theorem is that the replacement of operators with their eigenvalues 'ruins' certain operator identities, transforming them into algebraically inconsistent equations. The simplest example of this is

$$(\hat{\sigma}_x\hat{\sigma}_y)^2=-I\rightarrow(\sigma_x\sigma_y)^2=-1(?) \tag{2.6}$$

where $\hat{\sigma}_a$ and $\sigma_\alpha=\pm 1(\alpha=x,y)$ are the Pauli operators and their eigenvalues, respectively, and I is the unity operator. We note at once that the operator $\hat{\sigma}_x\hat{\sigma}_y=i\sigma_z$ is not Hermitian.

Interesting recent examples[24,25] include a few basic identities, each of which involves only mutually commuting combinations of operators; in order to reach a contradiction, all the identities must be multiplied so that the resulting 'contradictory relation' contains, as does (2.6), noncommuting operators (see Sec. 6.1 for more detail). These considerations reduce the essense of the paradox to a simple formula such as (2.6).

In the orthodox interpretation of QT, some of the properties of a particle, characterized by the operators $A\equiv A_\alpha$ and $A'\equiv A'_\alpha$, do not exist *a priori* (in the interference experiment, α and α' are different phase delays, with only one of them realized in a single sampling). This also applies to the properties B, B' of the other particle. Consequently, nonzero averages $E^{(1)}=\langle AB\rangle_\psi$, $E^{(2)}=\langle A'B\rangle_\psi$,... indicate correlations between things that do not exist! The situation is astonishing: the experiment involves several observers whose readings are uncertain, but have a predetermined correlation!

Such syllogisms have stimulated new attempts to reconcile the QT formalism with instinctive 'objective realism', using 'metaphysical' terms such as 'nonseparability', 'contrafactuality', 'contextuality', the notorious 'nonlocality', and varieties of these terms, none of which are rigorously defined.

One final point. It is widely thought that, in EPR-Bell type experiments, "...quantum mechanics predicts stronger correlations between particles than do local theories."[10] Its seems that this is in full agreement with the restriction $V_{\text{class}}\leqslant 0.5$ imposed on the interference visibility by classical theory. There are, however, other classical models that assure—as does QT—the full correlation $E=\pm 1$ (see Secs. 3.3, 4.4, and 5.4). Quantitative differences between QT and HVT are revealed only when account is taken of the specific dependence of the correlator $E(\alpha,\beta,...)$ on the phase parameters α, β,... of the operators A_α, B_β... (the BIs are not violated by all this). Thus, once again, 'nonlocality' and 'nonseparability' are seen to be undefined.

3. TWO-OBSERVER EXPERIMENTS

3.1. Two-channel interferometer

Figure 1a shows one of the simplest experimental arrangements[39] for the verification of BI. An essentially similar variant was proposed in Ref. 46 (see also Ref. 50) and implemented in Ref. 13.

A pump of frequency ω_0 and wave vector $\mathbf{k}_0$ illuminates a birefringent piezocrystal with quadratically nonlinear χ, which ensures that an extraordinary pump photon is split into a pair of ordinary photons with wave vectors $\mathbf{k}_a$ and $\mathbf{k}_b$. This is the so-called parametric scattering or spontaneous down-conversion effect.

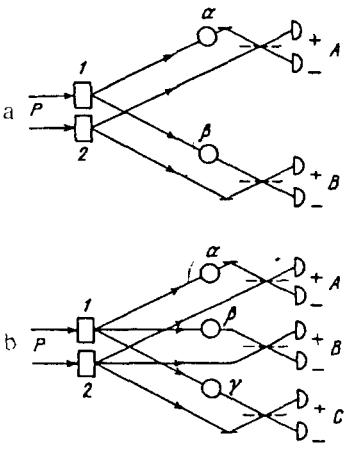

FIG. 1. Intensity interferometers with parametric sources of radiation and two (a) and three (c) observers. Correlated photons are produced simultaneously in the nonlinear elements 1 and 2 under the influence of the pump P and are sent to the observers A, B, (C) in two modes, one of which experiences a phase delay (circles). The modes are mixed in 50% beamsplitters (dashed lines) and are detected. In scheme (a) with zero resultant phase delay ($\varphi=\alpha+\beta=0$), both photons are synchronously deflected either upward (to detectors +) or downward (to detectors −); for $\varphi=\pi$, one travels upward and the other downward. In scheme (b) with $\varphi=\alpha+\beta+\gamma=0$, either all photons travel upward or one goes up and two go down. For $\varphi=\pi$, one or three go down.

The 'signal' photon $\mathbf{k}_a$ is distributed over two beams $\mathbf{k}_1^a$ and $\mathbf{k}_2^a$ by a mask that divides the exit face of the crystal into two regions 1 and 2. It is, of course, possible to use two separate crystals with a common coherent pump. This also happens to the second 'idle' photon whose frequency $\omega_b=\omega_0-\omega_a$ may not be equal to ω_a. The mask thus performs a *division of wavefront. Division of amplitude* is also possible and can be accomplished by means of beamsplitters[11,50] (see also Appendix III). Finally, we could exploit the axial symmetry of the radiated field:[13,46] a photon of a particular frequency ω_a (or ω_b) is allowed to enter a continuum of modes on the surface of a cone with an fixed scattering angle $\nu_a(\nu_b)$. Under these conditions, $(\mathbf{k}_{1,2}^a+\mathbf{k}_{1,2}^b)_\perp=0$. The aim of all these operations is the same: to divide the two beams a, b into four. In the case of polarization interferometers, modes 1, 2 must differ by the type of polarization. Similarly, we can produce $2j+1$ beam pairs that simulate correlated particles with spin j (Ref. 39).

Signal beams with the same frequency ω_a are mixed by a 50% beamsplitter and are directed on to two detectors D_+^a and D_-^a. In a preliminary step, a phase delay α is introduced into one of the beams. Similar elements are introduced into channel B: a phase delay β, a beamsplitter, and the detectors $D_\pm^b$.

The arrangement just described is essentially an intensity interferometer in which the nonlinear crystal is replaced with other sources. From the classical point of view,

the beamsplitters convert phase fluctuations into fluctuations in the intensity I that are recorded by the detectors. The result is that, when α or β vary slowly, each of the observed correlators $\langle I^a_\pm I^b_\pm \rangle$ oscillates harmonically with phase angle $\varphi=\alpha\pm\beta$ (Ref. 33 and 34). The sign depends on the source: the initial phases are correlated or anticorrelated (for further details, see Sec. 3.3 in which it is further shown that the oscillations can also be nonharmonic).

The interference visibility function V is determined by the relative noise level that is independent of φ. In QT it can reach unity, in HVT we find that $V\leqslant 1/\sqrt{2}$, and in classical stochastic electrodynamics the visibility V is limited by 1/2. All this applies to harmonic interference curves.

Suppose that the detectors operate as photon counters. The signal and idle photons are created almost simultaneously with a picosecond spread that is much smaller than the length T of the pulses produced by the detectors (for a photomultiplier, $T\sim 10^{-9}$ s). We choose the pump power to be low enough to ensure that, on average, we record pairs at intervals of, say, 10^{-6} s. The overlap probability for pulses due to adjacent (in time) photon pairs is then neglibly small.

If the quantum efficiency η of the detectors is 100%, they fire strictly in pairs: for example, a photon recorded by D^a_+ is necessarily accompanied by the simultaneous firing of D^b_+ or D^b_-. Complete correlation no longer occurs when $\eta<1$, and spurious events can be excluded by a coincidence system.

On the other hand, the detector counts in a given channel will be totally uncorrelated because photons are randomly directed either 'upward' or 'downward' by the beamsplitter. We shall follow Mermin[118] and, for convenience, connect green lamps to detectors D^a_+ and D^b_+ and red lamps to detectors D^a_- and D^b_-. Each detection of a pair will then be signalled by only two lamps: one in channel A and the other in channel B.

Let us now vary the phase in one of the channels, for example, α, and observe all the lamps. For a certain value of α, which we shall take as our origin ($\varphi=\alpha=0$), either both red or both green lamps will light up. If, however, we add π to α the lamps will light up out of step, i.e., red with green and vice versa.

We can now parametrize the picture by assigning to a function A the value $+1$ when detector D^a_+ fires and the value -1 when D^a_- fires. Similarly, we introduce the function $B=\pm 1$. We thus obtain two random 'point' processes A_i, B_i where i labels events that occur at random instants of time.

Finally, we define a third dichotomic (i.e., assuming only two discrete values) function $F_{\varphi i}\equiv A_{\alpha i}B_{\beta i}=\pm 1$. For $\varphi=0$, we always have $F_{0i}=1$, and for $\varphi=\pi$ we have $F_{\pi i}=-1$, i.e., we observe complete correlation or anticorrelation of the random sequences A_i and B_i, so that $F_{\varphi i}$ is then determined. Intermediate values of φ correspond to a reandom sequence $F_{\varphi i}$ with φ as parameter. We can measure its average value

$$E_\varphi\equiv\langle F_\varphi\rangle=\frac{1}{L}\sum_{i=1}^{L}F_{\varphi i},$$

$$F_{\varphi i}=A_{\alpha i}B_{\beta i},\quad \varphi=\alpha+\beta, \tag{3.1.1}$$

where L is the total number of pairs recorded in a particular interval of time. For 'sufficiently large' L, we obtain the 'frequency' distribution of the average.

We note that F_φ is a multichannel variable: to measure it, observers A and B must use communication links to exchange information (or to communicate it to a third person) and synchronize their clocks so as to establish the origin for the numeration of events, since each of the above random sequences forms a Poissonian random process. It is clear from the above discussion that signal transmission with superluminal velocity between observers A and B by means of quantum correlation is, of course, just as impossible as it is for clasical correlation.[139–143] However, this correlation can be used to to protect communication channels from eavesdropping[128–131] or from noise.[132–135]

We also emphasize that an observer who confines his attention to only one sequence, say, $A_{\alpha i}$, will not see any dependence on the phase α (or β) because in this case $P^+_A=P^-_A=1/2$. The other sequence has a similar distribution. At the same time, experiment and the quantum model suggest that

$$P^+_F(\varphi)=\cos^2(\varphi/2),\quad P^-_F(\varphi)=\sin^2(\varphi/2). \tag{3.1.2}$$

The joint distribution of observables A and B is

$$P^{++}_{AB}(\varphi)=P^{--}_{AB}(\varphi)=\frac{1}{2}\cos^2\frac{\varphi}{2},$$
$$P^{+-}_{AB}(\varphi)=P^{-+}_{AB}(\varphi)=\frac{1}{2}\sin^2\frac{\varphi}{2}. \tag{3.1.3}$$

Hence

$$E_\varphi=P^{++}_{AB}+P^{--}_{AB}-P^{+-}_{AB}P^{-+}_{AB}=\cos\varphi. \tag{3.1.4}$$

Significantly, in QT there is no joint distribution $P^{\pm\pm}_{AA'}$ for a pair of observables $A\equiv A_\alpha$ and $A'\equiv A'_\alpha$ when $\alpha\neq\alpha'$ (and similarly for B and B' when $\beta\neq\beta'$) because they do not commute: $[A,A']\propto\sin(\alpha-\alpha')$ [see also (III 22)] and cannot be measured simultaneously. It is precisely this that is the formal reason for the conflict between QT and HVT.

Total correlation is lost when 'random' coincidences due to overlap between neighboring pairs is taken into account. The right hand side of (3.1.4) then acquires an additional factor that can be interpreted as the visibility of the interference pattern:

$$E_\varphi=V\cos\varphi. \tag{3.1.5}$$

As the pump power is reduced and there is a corresponding reduction in the rate R at which biphotons are emitted, we have $V\to 1$ (see Sec. 3.3 for further details).

We now return to experiment. Let us establish in each channel two fixed values of the phase, α, α' and β, β', that differ by $\pi/2$, i.e.,

$$\alpha'-\alpha=\pi/2;\quad \beta'-\beta=\pi/2. \tag{3.1.6}$$

Next, we perform four series of measurements in succession, with the following combinations of phases:

$$\alpha,\beta;\quad \alpha,\beta';\quad \alpha',\beta;\quad \alpha',\beta'. \tag{3.1.7}$$

This yields the following four measured values of the multichannel observables:

$$F^{(1)}=AB,\quad F^{(2)}=AB',\qquad F^{(3)}=A'B,\quad F^{(4)}=A'B'. \tag{3.1.8}$$

We now define the combination

$$S\equiv\tfrac{1}{2}(F^{(1)}+F^{(2)}+F^{(3)}-F^{(4)}), \tag{3.1.9}$$

which we shall call the Bell observable. The data are then averaged:

$$\langle S\rangle_{\exp}=\tfrac{1}{2}(E^{(1)}+E^{(2)}+E^{(3)}-E^{(4)})=\tfrac{1}{2}\langle AB+AB'+A'B-A'B'\rangle_{\exp}. \tag{3.1.10}$$

This result should be close to the QT prediction that follows from (3.1.4) and (3.1.6):

$$\langle S\rangle_{\psi}=\frac{1}{2}\left[\cos\varphi_1+2\cos\left(\varphi_1+\frac{\pi}{2}\right)-\cos(\varphi_1+\pi)\right]=\sqrt{2}\cos\left(\varphi_1+\frac{\pi}{4}\right), \tag{3.1.11}$$

where $\varphi_1=\alpha+\beta$ and $\langle\ldots\rangle_{\psi}$ represents averaging over the quantum state characterized by the vector $|\psi\rangle$ which will be defined later [see (3.4.1)]. The Bell observable is a maximum when $\varphi_1=-\pi/4$:

$$\langle S(A,A',B,B')\rangle_{\psi,\max}=\sqrt{2}. \tag{3.1.12}$$

It may be shown that the value $\sqrt{2}$ cannot be exceeded [see (III25)] whatever the state $|\psi\rangle$ in which the system finds itself.[87–89]

It is interesting to note that $\langle S\rangle_{\psi}$ depends only on the sum $\alpha+\beta$, leaving one of the component phases (say, α) free for the chosen φ_1. We also note that all four terms in (10) give the same positive contribution (or the same negative contribution for $\varphi_1=3\pi/4$) to $\langle S\rangle_{\psi}$, which is equal to $1/\sqrt{2}$ when $\langle S\rangle_{\psi}$ reaches its extremal values.

The same experiment can be considered within the framework of HVT under certain 'natural' and fully 'reasonable' assumptions, namely, that the variables A, A', B, B' have *a priori* values in the range between -1 and $+1$ that are specified by a certain nonnegative joint distribution function, and that the measuring devices do not influence one another (the assumptions are discussed in greater detail in, for example, Refs. 41, 44, and 45). This leads to

$$-1\leqslant\langle S\rangle_{\rho}\leqslant 1. \tag{3.1.13}$$

The conflict between (3.1.12) and (3.1.13) can serve as a possible criterion for choosing between the two theories.

We emphasize that the general restriction defined by (3.1.13) must be met by all possible special classical models describing experimental procedures analogous to those that we have described. Some of them are examined in this review.

We also note that inequality (3.1.13) was established for the *average* Bell variable. On the other hand, in individual combinations such as (3.1.9), we have $S=0$, $\pm1,\pm2$ in both classical and quantum theories.

3.2. Bell inequalities for two observers

Let us now try to explain the picture described above, i.e., the flashing of the red and green lamps, in terms of very general propositions. Let us forget about Maxwell equations, the Schrödinger equation, photons, interference, and so on. We follow Laplace: if two green lamps have flashed at time t_i, i.e., the event "$A_i=1$, $B_i=1$" has taken place, then it must have been due to some prior causes, i.e., a 'confluence of circumstances'. Let λ represent the set of parameters that uniquely lead to this event. We cannot measure all these parameters, either in principle or because of their number (say, 10^{23}). We shall therefore refer to them as "hidden parameters." However, we may assign to them a specific distribution function $\rho(\lambda)$ where

$$\int\rho(\lambda)\mathrm{d}\lambda=1, \tag{3.2.1}$$

in which the integral is evaluated over the entire set λ of of the parameter values. The flash times t_i, connected with the dynamics $\lambda(t)$, are also uniquely defined by $\lambda_i\equiv\lambda(t_i)$.

Thus, we believe that there are determined (single-valued) dichotomic functions $A_{\alpha i}=A(\alpha,\lambda_i)$, $B_{\beta,i}=B(\beta,\lambda_i)$ in which the 'revealed' parameters α, β and the 'locality' of the measurements are separately indicated: A is independent of β and B of α. The phase-controlling observers can thus be as far as desired from one another and from the light source, i.e., from the carrier of the hidden parameters. Consequently, there is also a third determined function

$$F_{\alpha\beta i}\equiv A_{\alpha i}B_{\beta i}. \tag{3.2.2}$$

The four series of measurements described in the last Section are determined by the functions

$$A(\lambda_i^{(m)}),\ A'(\lambda_i^{(m)}),\ B(\lambda_i^{(m)}),\ B'(\lambda_i^{(m)}); \tag{3.2.3}$$

where $\lambda_i^{(m)}=\lambda(t_i^{(m)})$ and $t_i^{(m)}$ is the time of the i-th event in the m-th series.

Each of the set of functions $\lambda(t)$ varies appreciably on the atomic time scale, so that they can be looked upon as random processes in which the values of λ at times $t_i^{(m)}$, $t_{i+1}^{(m)}$, $t_i^{(m\prime)}$ are independent. We now adopt the ergodic theorem and replace time averages by ensemble averages with weight $\rho(\lambda)\equiv\rho_{\lambda}$. Discarding the indices on the functions in (3.2.3), we obtain

$$\langle S\rangle_{\rho}=\int S_{\lambda}\rho_{\lambda}\mathrm{d}\lambda, \tag{3.2.4}$$

where in accordance with (3.1.9)

$$S_{\lambda}\equiv\tfrac{1}{2}(A_{\lambda}B_{\lambda}+A_{\lambda}B'_{\lambda}+A'_{\lambda}B_{\lambda}-A'_{\lambda}B'_{\lambda}). \tag{3.2.5}$$

It is readily verified that, like A_λ and B_λ, the function S_λ is dichotomic. This can be done by writing

$$S_\lambda=\tfrac{1}{2}[A_\lambda(B_\lambda+B'_\lambda)+A'_\lambda(B_\lambda-B'_\lambda)]=\pm 1. \qquad (3.2.6)$$

If $B_\lambda = B'_\lambda$ then $S_\lambda=A_\lambda B_\lambda=\pm 1$, and if $B_\lambda = -B'_\lambda$, then $S_\lambda = A'_\lambda B_\lambda = \pm 1$. This does not necessarily ensure that the phase restriction (3.1.6) is met. The significant point is that all the functions in (3.2.6) must have the single (equal) argument λ, since otherwise (3.2.6) may not be satisfied. We emphasize this particularly because this point is not usually brought out in the literature. The indices λ will often be discarded in the discussion given below.

In classical probability theory, and according to common sense, $S_\lambda \geqslant 0$, so that it follows from (3.2.1), (3.2.4), and (3.2.6) that

$$|\langle S\rangle_\rho| \leqslant \int |S_\lambda|\rho_\lambda \mathrm{d}\lambda = 1. \qquad (3.2.7)$$

Thus, finally,

$$-1 \leqslant \langle S\rangle_\rho \leqslant +1. \qquad (3.2.8)$$

This is one of tHe simplest variants of BI. Its other modifications and generalizations are given, for example, in Refs. 15, 40, 41, 45, and 89. Here we merely mention the fact that the dichotomy of the functions A, A', B, B' is not a necessary condition for the validity of (3.2.8). It is sufficient to demand that

$$|A|\leqslant 1,\ \ |B|\leqslant 1,\ \ |A'|\leqslant 1,\ \ |B'|\leqslant 1, \qquad (3.2.9)$$

i.e., the measurements can also be performed in the continuous spectrum of values, but necessarily normalized.

Let us now return to the quantum interpretation of the above experiment.

We note that, for complete correlation, when

$$\langle AB\rangle = \langle AB'\rangle = \langle A'B\rangle = \langle A'B'\rangle = \pm 1, \qquad (3.2.10)$$

the BI is still satisfied—though only just [see (3.1.11) with $\varphi_1=0$ or π]—in contrast to the three-channel version of the experiment in which it is precisely for complete correlation that the violation of the BI is quantitatively at its maximum (see Sec. 4). The logic of the EPR program does not therefore apply to the possible violation of (3.2.8). We recall in this connection that the original EPR program provided for the augmentation of quantum mechanics with hidden parameters[1,15] as well as experiments ensuring the complete correlation of measurements, since the principal intention was to eradicate the *statisticity* and to verify determinism ('God does not play dice'). This means that, strictly speaking, the statistical character of the violation of BI in two-observer experiments has no relevance for the EPR paradox. In other words, when the experiment described in Sec. 3.1 is analyzed, we have to distinguish two situations, namely, complete correlation ($E=\pm 1$) to which EPR logic applies and incomplete correlation for which $E=\pm 1/\sqrt{2}$ and (3.2.8) is violated in QT.

3.3. Classical stochastic model

We must now try to describe the experiment illustrated in Fig. 1a in the language of classical statistical theory, using the model of interfering waves with fluctuating phases. This type of experiment is readily performed in the radiofrequency range.

Suppose that we use 'single-mode' detectors (cf., for example, Ref. 154): their time constant T and transverse size R of the aperture should be much smaller than the corresponding coherence scales of the radiation incident upon them. The radiation is assumed to be quasimonochromatic and quasiplane: $T \ll \tau_{\rm coh} \propto 2\pi/\Delta\omega$, $R \ll \rho_{\rm coh}$. Such detectors produce a signal $i(t)$ that is proportional to the instantaneous intensity, i.e., $i(t)=\eta n(t)=\eta|a(t)|^2$ where η is the detector efficiency, proportional to R^2T, and $a(t)$ is a slowly-varying (on the scale of scale $\tau_{\rm coh}$ field amplitude in dimensionless units (in which energy flux is $\hbar\omega|a|^2\Delta\omega/2\pi$ (Ref. 154).

We must now elucidate the effect of the phase delay α and the 50% beamsplitter that mixes two spatial modes in one channel (cf. Fig. 1a). If the modes differ only by the type of polarization, we can use a Nicol prism as the beamsplitter. Suppose that $a_k=|a_k|\exp(-ix_k)$ are the cmplex amplitudes at the input ($k=1,2$).

The output amplitudes can then be written in the form

$$a_\pm = (\pm a_1 e^{i\alpha/2} + a_2 e^{-i\alpha/2})/\sqrt{2}. \qquad (3.3.1)$$

The common phase factor is ignored because it does not appear in the output intensities:

$$n^a_\pm \equiv |a_\pm|^2 = [n_a \pm 2|a_1a_2|\cos(x+\alpha)]/2, \qquad (3.3.2)$$

where

$$n_a \equiv n^a_1 + n^a_2 \equiv |a_1|^2 + |a_2|^2, \quad x(t)\equiv x_2 - x_1$$

are the total intensity in channel A and the phase difference that is a slowly-varying function of time, respectively.

We note that $|a_1+a_2|^2 \geqslant 0$ leads to

$$2|a_1a_2| \leqslant |a_1|^2 + |a_2|^2 \equiv n_a. \qquad (3.3.3)$$

The transformation given by (3.3.1) applies to a type of unitary SU(2) transformations and conserves energy: $n^a_+ + n^a_- = n^a_1 + n^a_2 \equiv n_a$ (see, for example, Refs. 33, 34, and 155).

According to (3.3.2), the beamsplitter transforms fluctuations in the phase difference $x(t)$ into fluctuations in intensity $n^a_\pm(t)$. For $T \ll \tau_{\rm coh}$, (3.3.2) describes a stationary *amplitude interference*, i.e., a harmonic dependence of n^a_+, n^a_- on the parameter α. If, on the other hand, the averaging onterval is $T > \tau_{\rm coh}$, the average intensities do not depend on α: $\langle n^a_\pm\rangle = \langle n_a\rangle/2$.

A dependence analogous to (3.3.2) determines the output intensities $n^b_\pm$ in channel B in terms of the input amplitudes and phases $b_k=|b_k|\exp(-iy_k)$.

Consider the correlation between the intensities of two arbitrary output modes in channels A and B; for example, according to (3.3.2)

$$\langle n^a_+ n^b_+\rangle = \tfrac{1}{4}\langle n_a n_b\rangle + \langle |a_1a_2b_1b_2|\cos(x+\alpha) \times \cos(y+\beta)\rangle. \qquad (3.3.4)$$

If the input intensities do not fluctuates, or fluctuate independently of the phases, the second term is proportional to the sum

$$\langle\cos(x(t)+y(t)+\alpha+\beta)\rangle+\langle\cos(x(t)-y(t)+\alpha-\beta)\rangle. \tag{3.3.5}$$

Consequently, stationary intensity interference can be observed only in the following two cases:

$$x(t)\pm y(t)=\text{const}, \tag{3.3.6a}$$

i.e., when the phases of the initial waves are correlated or anticorrelated (see, for example, the review given in Ref. 33). We note that the paired mixing of modes can be achieved directly without using beamsplitters, e.g., by employing photosensitive surface detectors. Two-mode and even multimode interference are also possible.[21,33,34,156,157]

In quantum theory, conditions (3.3.6a) correspond to nonzero values of the following two types of correlator:[33,34]

$$G_+\equiv\langle a_1a_2^+b_1b_2^+\rangle_\psi\neq 0,$$

or (3.3.6b)

$$G_-\equiv\langle a_1a_2^+b_1^+b_2\rangle_\psi\neq 0,$$

where a^+, b^+ and a, b are the photon creation and annihilation operators. When spin experiments are described, the correlators $G_\pm$ are replaced with $\langle\sigma_-^a\sigma_\mp^b\rangle$ (see Appendix I).

The case $G_-\neq 0$ or $x-y=\text{const}$ constitutes a response to the well-known Brown–Twiss[42] intensity interference (see also Ref. 43, p. 106). The pair of light sources can then be, for example, two stars.

Anticorrelation of phases $G_=\neq 0$ or $x+y=\text{const}$ arises when 'parametric' noise or two-photon atomic transitions are employed. These two cases have a number of common features.[65,66]

It is thus clear that there are two main types of intensity interference with simple classical explanation: correlated or anticorrelated phase fluctuations, transformed into additional amplitude fluctuations. We note, that in contrast to the above four-mode scheme (see Fig. 1a), the correlation or anticorrelation of phases in the form described by (3.3.6a) and (3.3.6b) is not essential in two-mode interference.[33,34] This special case can be classified as the third basic type of intensity interference. For parametric generators, the phase $x+y=x_2-x_1+y_2-y_1$ is determined by the constant phase difference between the pump waves (see Fig. 1a). Let us put $x+y\equiv 0$, so that for constant input intensities n_a, n_b, we have from (3.3.4)

$$\langle n_+^a n_+^b\rangle=\tfrac{1}{4}n_an_b(1+V\cos\varphi), \tag{3.3.7}$$

where $\varphi\equiv\alpha+\beta$ and the visibility is given by

$$V=\frac{2|a_1a_2b_1b_2|}{n_an_b}\leqslant\frac{1}{2}. \tag{3.3.8}$$

The last inequality follows from (3.3.3). It is thus evident that in the classical model with constant input intensities, the visibility will not exceed 1/2. When initial Gaussian intensity fluctuations are taken into account, this limit falls to 1/3. A further reduction in V occurs when the condition $T\ll\tau_{\text{coh}}$ is violated (see Sec. 3.5). An analogous model gives the following expression for three-channel intensity interference (Fig. 1b) when (3.3.3) is taken into account for constant input intensities n_a, n_b, n_c:

$$V_3=\frac{2|a_1a_2b_1b_2c_1c_2|}{n_an_bn_c}\leqslant\frac{1}{4}, \tag{3.3.9}$$

whereas for N channels

$$V_N\leqslant 1/2^{N-1}. \tag{3.3.10}$$

Let us now return to the scheme of Fig. 1a and consider the following differences [in quantum theory, these observables can also be expressed in terms of phase difference operators; see (I19)]:

$$\Delta n_a\equiv n_+^a-n_-^a=a_1a_2^*e^{i\alpha}+\text{c.c.}=2|a_1a_2|\cos(x+\alpha),$$

$$\Delta n_b\equiv n_+^b-n_-^b=b_1b_2^*e^{i\beta}+\text{c.c.}=2|b_1b_2|\cos(y+\beta). \tag{3.3.11}$$

When the parametric condition $x+y=0$ and the possibility of phase-independent intensity fluctuations are taken into account, their correlation normalized to $\langle n_an_b\rangle$ takes the form

$$E_\varphi\equiv\frac{\langle\Delta n_a\Delta n_b\rangle}{\langle n_an_b\rangle}=V\cos\varphi, \tag{3.3.12}$$

$$V=2\langle|a_1a_2b_1b_2|\rangle/\langle n_an_b\rangle. \tag{3.3.13a}$$

For constant n_a and n_b, this normalization ensures the validity of the Bell theorem because, in any realization, the modulus of the measured relative variables does not exceed unity [see (3.2.9)]:

$$\frac{|\Delta n_a|}{n_a}\equiv\frac{|n_+^a-n_-^a|}{n_+^a+n_-^a}\leqslant 1,$$

$$\frac{|\Delta n_b|}{n_b}\equiv\frac{|n_+^b-n_-^b|}{n_+^b+n_-^b}\leqslant 1. \tag{3.3.14}$$

Thus, in the 'classical' version of the experiment designed to verify BI, the intensities of the parametrically generated waves must be maintained constant in all the series of measurements, which is readily achieved when parametric generators are in fact used.

In quantum theory, (3.3.13a) is replaced with

$$V=\frac{2\,\text{Re}\langle a_1a_2^+b_1b_2^+\rangle_\psi}{\langle n_an_b\rangle_\psi} \tag{3.3.13b}$$

and, as will be shown in Sec. 3.4, the visibility can become equal to unity. Moreover, there are states $|\psi\rangle$ for which $V_N=1$ for arbitrarily N (see Sec. 5).

Thus, an increase in the number N of channels is accompanied by an exponential increase in the relative difference between the classical and quantum predictions of visibility (of the form 2^{N-1}) and becomes noticeable from $N=2$ onwards. This results is conveniently different from the Bell theorem prediction (5.1.8) in which the relative increase in the discrepancy is only $2^{(N-1)}/2$. However, in

this case, we are confining our attention to a particular model of the experiment, in contrast to the general formulation of the problem in the Bell theorem.

It is interesting to note that, if the correlation between Δn_a and Δn_b is normalized to the variance of fluctuations in the difference $\langle \Delta n^2 - \langle \Delta n \rangle^2 = \langle \Delta n^2 \rangle$, i.e., if we consider the usual correlation coefficient

$$\Gamma_\varphi = \frac{\langle \Delta n_a \Delta n_b \rangle}{(\langle \Delta n_a^2 \rangle \langle \Delta n_b^2 \rangle)^{1/2}} = \cos\varphi, \tag{3.3.15}$$

the classical and quantum interpretations become identical for two-channel interferometers. The Bell theorem condition (3.3.9) is not of course satisfied in this interpretation because the measured relative quantities $\Delta n_a \langle \Delta n_a^2 \rangle^{1/2}$ and $\Delta n_b / \langle \Delta n_b^2 \rangle^{1/2}$ can exceed unity. This was first pointed out by Barut and Meystre[93] who analyzed the behavior of two particles with anticorrelated angular momenta. Their scheme is described in greater detail in Appendix II. Here we merely note that, in this case, the visibility is limited to 1/3, which depends on the number of equally probable projections of the three-dimensional angular momentum vector.

The above classical models describe experiments with analog detectors that produce readings with a continuous spectrum of values. The interference structure then takes the form $\cos\varphi$ as in QT, but the maximum classical correlation does not exceed 1/2 in optical experiments and 1/3 in spin experiments, which is different from QT which allows complete correlation, i.e., correlation equal to unity. This difference is due to the use of continuous variables whereas photon counters give $n=0$ or 1, so that $\Delta n_{a,b}$ are dichotomic variables: $\Delta n_{a,b} = \pm 1$ (in the case of one-photon states in each channel).

Complete correlation $E = \pm 1$ can also be obtained classical theory by using dichotomic observables with a discrete spectrum (see, for example, the spin models described in Refs, 3 and 15). We now turn to a discrete wave model.

Let us suppose that we periodically record the readings of four detectors (Fig. 1a) in steps of $\Delta t \gg \tau_{\rm coh} \gg T$. This gives a sequence of random numbers $n_\pm^{a,b}(t_i)$, $i=1,2,\ldots,L$ where $L\Delta t$ is the total duration of the sequence. We now use the two difference files $\Delta n_{a,b}(t_i)$ to form three dichotomic "sign" sequences

$$A_i \equiv \text{sign}\, \Delta n_a(t_i), \quad B_i \equiv \text{sign}\, \Delta n_b(t_i),$$

$$F_i \equiv A_i B_i. \tag{3.3.16}$$

Thus, $A_i = +1$ symbolizes an event in which the phtocurrent from "detector D_+^a is greater than that from D_-^a."

We shall take as our source two parametric generators with a common coherent pump, so that the condition $x(t) + y(t) = 0$ is satisfied. This means that the instantaneous frequencies $\omega(t) = \bar{\omega} + {\rm d}x/{\rm d}t$ of signal (a) and the idle (b) waves always drift in opposite directions. We note that $\omega_a = \omega_b$ is not a necessary condition: for stationary interference (without beats) we need only have $\omega_1^a = \omega_2^a$ and $\omega_1^b = \omega_2^b$.

According to (3.3.11) and (3.3.16), we have

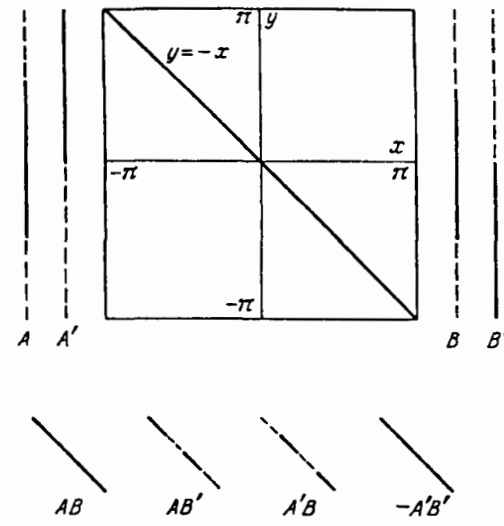

FIG. 2. Diagrams illustrating the relation between hidden parameters (random phases x and y) and the sampled dichotomic observables in the classical model. Parametric limitation ensures that the phases correspond to the bisector $x=-y$. Thick solid lines correspond to observables equal to $+1$ and and dashed lines to -1; $A=A_0$, $A'=A_{\pi/2}$, $B=B_0$, $B'=B_{\pi/2}$.

$$A_{\alpha i} = \text{sign}\cos(\alpha + x_i), \tag{3.3.17}$$

$$B_{\beta i} = \text{sign}\cos(\beta - x_i),$$

$$F_{\alpha\beta i} = \text{sign}[\cos(\alpha+\beta) + \cos(\alpha - \beta + 2x_i)],$$

where $x_i \equiv x(t_i)$. Consequently, when $\varphi \equiv \alpha + \beta = 0$ or π, we have complete correlation (or anticorrelation) $F_{\alpha\beta i} = +1$ (or -1) as in QT.

From the classical standpoint, the function $x(t)$ is completely determined by processes in the parametric generator (possibly at the atomic level), i.e., there is a single-valued relation $x_i = x[\lambda(t_i)]$ where $\lambda(t)$ is the set of all parameters influencing the phase. It is thus natural to regard the phase itself as a 'hidden' parameter over which the average is evaluated. The relation between the random phases $x(t)$ and $y(t)$ on the one hand, and the recorded values of the observables is clearly illustrated in Fig. 2. Suppose that $x(t)$ has the uniform distribution $\rho(x) = 1/2\pi$ in the interval $-\pi$, π, in which case

$$E_\varphi = \frac{1}{2\pi} \int_{-\pi}^{\pi} F_{\alpha\beta x} {\rm d}x$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{sign}[\cos\varphi + \cos(2x)]\,{\rm d}x$$

$$= 1 - 2\frac{|\varphi|}{\pi} \tag{3.3.18}$$

for $|\varphi| \leqslant \pi$. A derivation of this is given in Appendix IV and a graph of E_φ is reproduced in Fig. 3b.

The Bell observable now assumes the form

$$\langle S \rangle_\rho = 1 + [(|\alpha+\beta| + |\alpha+\beta'| + |\alpha'+\beta| - |\alpha'+\beta'|)]/\pi. \tag{3.3.19}$$

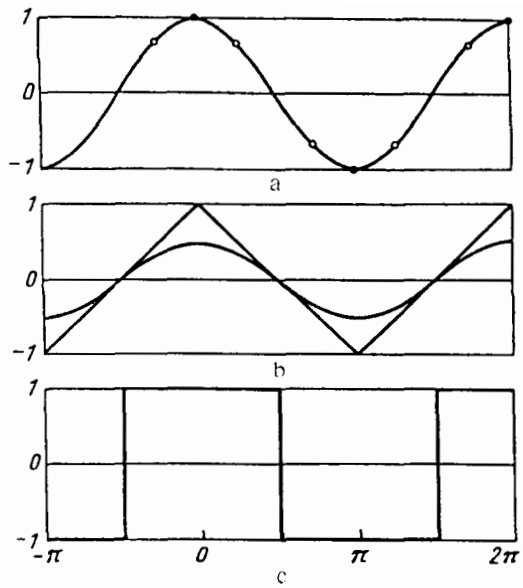

FIG. 3. Interference curves, i.e., graphs of the correlator E as a function of the sum of channel phase shifts $\varphi=\alpha+\ldots$ (a,b) and one of these shifts (α_n) (c) for the following models: a—+quantum model, number of observers $N\geqslant 2$ arbitrary, open and full points represent optimum values of the phases for even and odd N, b—classical model with uniformly distributed random phases for $N=2$ (broken line) and $N=3$ (mutually inverted segments of parabolas), c—classical model with discrete phase distribution (random signs of amplitudes), $N\geqslant 2$ arbitrary.

When we use the combination of phases given by (3.1.11) with $\varphi_1=\pi/4$, which ensures the maximum violation of BI in quantum experiments, we have

$$E^{(1)}=E^{(2)}=E^{(3)}=-E^{(4)}=1/2, \tag{3.3.20}$$

so that $\langle S\rangle_\rho=1$ and the BI is not violated. This is not unexpected because the system that we have described is a particular implementation of the universal Bell model. On the other hand, the BI may not be satisfied in individual realizations because, for example, A and A' are recorded at different times with different values of the 'hidden' parameters $x(t_i)$ [see also (3.2.6) and the following text].

It is now a relatively simple matter to use (3.3.16) and (3.3.17) to find the joint probability distributions $P^{\pm\pm}_{AB}$, $P^{\pm\pm}_{A,A'}$ that are analogous to (3.1.3). However, the existence of observables A and A' that are incompatible in QT is a specific feature of HVT. We recall that the distributions $P^{\pm\pm}_{AA'}$ do not exist in QT for $\alpha\neq\alpha'$.

Thus, although the above model provides complete correlation (as in QT), the 'rectification' of the harmonic interference curve $E_\varphi=\cos\varphi$ into a sawtooth curve (see Fig. 3) ensures that the model does not violate BI. We note that, according to Appendix I, this scheme is isomorphic with the model of two classical particles with spin, discussed by Bell[3] (see also Ref. 15), which also leads to the linear relation (3.3.18).

3.4. Quantum theory of two-photon interference

In the experimental scheme of Fig. 1a, the source is assumed to be a parametric amplifier-converter that incorporates piezocrystals. It spontaneously produces photon pairs ('biphotons') with correlated times of creation (or energies),[65] and also correlated directions of emission. Parametric converters are the most effective sources of two-photon light.

It will be shown in Appendix III that two such converters with a common coherent pump will frequently prepare an initial four-mode field in the state

$$|\psi\rangle=\frac{1}{\sqrt{2}}(a_1b_1+a_2b_2)^+|0\rangle$$

$$\equiv\frac{1}{\sqrt{2}}(|10\rangle_a|10\rangle_b+|01\rangle_a|01\rangle_b), \tag{3.4.1}$$

where $a^+_{1,2}$, $b^+_{1,2}$ are photon creation operartors in the signal and idle nodes, and $|0\rangle\equiv|0\rangle^a_1|0\rangle^b_1|0\rangle^a_2|0\rangle^b_2$ represents the vacuum state. We interpret (3.4.1) as follows: the signal (idle) photon can be created with equal probability in one of the two crystals, i.e., it belongs to two modes a_1 and a_2 at the same time, but the associated idle (signal) photon is necessarily created in the same crystal. When the pump power is low enough, so that not more than one pair of photons is transmitted during the time T of a single measurement, we may consider that the state (3.4.1) corresponds to each single measurements. Such nonfactorizable states, consisting of two or a large number of particles, are called entangled. Many of the quantum mechanical paradoxes are associated with these states.

The effect of the phase delays α, β and of the beamsplitters (Fig. 1a) on the initial field will now be described in the Heisenberg representations, i.e., with the aid of (3.3.1). The photon number operators for the two output modes of channel A then assume the form [cf. (3.3.2)]

$$n^a_\pm\equiv a^+_\pm a_\pm=\tfrac{1}{2}[n^a_1+n^a_2\pm(\sigma^a_-e^{i\alpha}+\sigma^a_+e^{-i\alpha})], \tag{3.4.2}$$

where

$$\sigma^a_-\equiv a_1a_2^+,\sigma^a_+\equiv a_1^+a_2=(\sigma^a_-)^+. \tag{3.4.3}$$

Hence we find that the operator corresponding to the observed event, i.e., a flash of the green or red lamp in channel A, is

$$A_\alpha\equiv n^a_+-n^a_-=\sigma^a_-e^{i\alpha}+\text{h.c.}, \tag{3.4.4}$$

where h. c. represent the Hermitian conjugate.

The operators $n^b_\pm$, B_β can be defined by analogy.

We emphasize that our *quantum* observables have precisely the same 'locality' property as the *classical* variables: A_α is independent of β and B_β is independent of α. The Heisenberg representation that we are using enables us to treat the locality of QT as an obvious fact. Actually, the state vector (3.4.1) at the interferometer input determines the statistical properties of the light source, i.e., a pair of parametric converters, until the photons depart in opposite directions. Their subsequent state, i.e., propagation in linear devices, can be adequately described by classical laws.[155]

The time delays in interferometer channels did not arise in (3.4.2) and (3.4.4) because of the use of monochromatic modes. Allowance for the finite width of the

spectra, i.e., multimode description of two-photon interference, introduces nothing significant into the question of locality.[156]

The 'nonseparability' of the quantum picture in such experiments is often referred to, meaning the nonfactorizability of the state vector of the particle pair whose separation has become large. However, this concept arises only in the Schrödinger approach that involves the wave function of a pair that has already separated, i.e., it takes the propagation effect into account. The equivalent Heisenberg description clearly shows that QT is just as local as, for example, classical stochastic electrodynamics in which interesting correlations, analogous to quantum correlations, are also possible (see Sections 3.3 and 3.6, and also Appendix II).

We now return to our main discussion and use (3.4.4) to find the multichannel operator

$$F_\varphi \equiv A_\alpha B_\beta = \sigma_-^a \sigma_-^b e^{i\varphi} + \text{h.c.}, \quad \varphi = \alpha + \beta, \tag{3.4.5}$$

where we have discarded the terms $\sigma_\mp^a \sigma_\pm^b \exp i(\pm\alpha \mp \beta)$ because, when applied to the vector (3.4.1), they yield zero.

We note that, in general, $A \equiv A_\alpha$ and $A' \equiv A_{\alpha'}$, $B \equiv B_\beta$ and $B' \equiv B_{\beta'}$, and also F_φ and $F_{\varphi'}$ do not commute (see Appendix III).

Consequently, the corresponding observables cannot be measured in a single sampling. This is not difficult to understand: for example, the phase α cannot have two values at the same time.

It is readily verified that (3.4.1) is the eigenvector with eigenvalues ± 1 of the following operators

$$A_\alpha^2 = B_\beta^2 = F_\varphi^2 = I, \tag{3.4.6}$$

$$(AA')^2 = (BB')^2 = (F_{\alpha\beta} F_{\alpha'\beta'})^2 = -I, \tag{3.4.7}$$

$$F_0 |\psi\rangle = |\psi\rangle, \tag{3.4.8}$$

$$F_\pi |\psi\rangle = -|\psi\rangle, \tag{3.4.9}$$

where I is the unity operator, $\alpha' = \alpha \pm \pi/2$, $\beta' = \beta \pm \pi/2$, $F \equiv \sigma_x^a \sigma_x^b$, $F_\pi \equiv -\sigma_x^a \sigma_x^b$, and σ_x is the Pauli matrix (see also Appendix I). In the identities (3.4.6) and (3.4.7), we have used the more general expressions given by (3.4.4) instead of (3.4.5). We can now understand why the observable F_0 does not fluctuate in the experiment. It always assumes the value $+1$ in the state (3.4.1), i.e., we have complete correlation. Similarly, F_π is always equal to -1 (anticorrelation).

The negative eigenvalues of the squares of the operators in (3.4.7) are due to the following property of Pauli matrices[158] (see also Appendix I):

$$(\sigma_x \sigma_y)^2 = (i\sigma_z)^2 = -1. \tag{3.4.10}$$

Although operators such as $\sigma_x \sigma_y$ are formally possible, they cannot describe single measurements and are non-Hermitian (like AA' in the general case):

$$(\sigma_x \sigma_y)^+ = \sigma_y \sigma_x = -\sigma_x \sigma_y. \tag{3.4.11}$$

In the state described by (3.4.1), the correlator is $\langle \sigma_-^a \sigma_-^b \rangle = 1/2$, so that according to (5)

$$\langle F_\varphi \rangle_\psi = \cos\varphi. \tag{3.4.12}$$

When (3.4.6) and (3.4.12) are taken into account, the variance of fluctuations in the multichannel obervable is given by

$$\langle \Delta F_\varphi^2 \rangle \equiv \langle (F_\varphi - \langle F_\varphi \rangle)^2 \rangle = 1 - \langle F_\varphi \rangle^2 = \sin^2\varphi. \tag{3.4.13}$$

As before, we define the Bell operator for even values of the phases:

$$S \equiv \tfrac{1}{2}(F^{(1)} + F^{(2)} + F^{(3)} - F^{(4)})$$
$$= \tfrac{1}{2}\sigma_-^a \sigma_-^b (e^{i\varphi_1} + e^{i\varphi_2} + e^{i\varphi_3} - e^{i\varphi_4}) + \text{h.c.} \tag{3.4.14}$$

We note that, according to this definition, S is the sum of four noncommuting operators, which has a corresponding observable that must be measured by four samplings with different phases φ.

The average of (3.4.14) is

$$\langle S \rangle_\psi = \tfrac{1}{2}(\cos\varphi_1 + \cos\varphi_2 + \cos\varphi_3 - \cos\varphi_4). \tag{3.4.15}$$

According to (3.4.13)–(3.4.15), the variance of the Bell observable is given by

$$\langle \Delta S^2 \rangle_\psi = \frac{1}{4} \sum_{m=1}^{4} \langle \Delta F^{(m)^2} \rangle = \frac{1}{4} \sum_{m=1}^{4} \sin^2\varphi_m. \tag{3.4.16}$$

Here we have simply taken the sum of variances of fluctuations in each series of measurements because taking a 'quartet' of experimental results does not imply any particular algorithm and can be a purely fortuitous choice (the only restriction is that we have to avoid repetition, i.e., using the same outcomes of samplings in different 'quartets'). Thus, there is no coupling between $F^{(m)}$ and $F^{(m')}$ for $m \neq m'$.

The combination of phases

$$-\varphi_1 = \varphi_2 = \varphi_3 = \varphi_4/3 = \pi/4, \tag{3.4.17}$$

that originally appeared in (3.4.15) gives $\langle S \rangle_\psi = \sqrt{2}$, which agrees with the exact upper limit for the norm of the operator S (Refs. 87–89) and is greater by 41% than the maximum classical value (unity). According to (3.4.16), we then have $\langle \Delta S^2 \rangle = 1/2$.

We recall that, in classical statistical models, the formula given by (3.4.12) is replaced with $V\cos\varphi$ where $V \leqslant 1/2$ [see (3.3.12)] or $1 - (2|\varphi|/\pi)$ [see (3.3.18)], which explains their agreement with BI.

We also note that, in discrete HVT models, all four variables A, A', B, B' can only assume the values ± 1 (they are dichotomic) and the Bell inequality $|\langle S \rangle| \leqslant 1$ follows from the classical condition $S_\lambda = \pm 1$ [see (3.2.6), (3.2.7) and (3.3.16)]. However, the observables A, A', B, B' have the same spectrum in QT. The question then is: why is the equation $S = \pm I$ not satisfied in QT? It turns out that it is precisely the nonzero commutators $[A,A']$ and $[B,B']$ that increase the norm of the operator S beyond its classical limit (unity)[87–89] (see Appendix III).

The fact that the operators do not commute means that they do not have common eigenvectors and, consequently, there are no joint distributions, e.g., we cannot

define $P_{AA'}$ for $\alpha \neq \alpha'$ in QT. On the other hand, in the classical approach, the existence of the distribution is postulated (explicitly or implicitly), which imposes an additional restriction on the Bell observable S (Ref. 107).

On the other hand, the operators A_α and B_β commute and their joint distributions (3.1.3) can be evaluated with the help of general QT recipes (see Appendix III).

It is shown in Appendix I that the fact that the operators A_0 and $A_{\pi/2}$ do not commute is analogous to the noncommutation of the Pauli matrices σ_x and σ_y that correspond to the position and momentum operators q_k, p_k of the two oscillators. A further useful interpretation is presented in Appendix I and shows that A_0 and $A_{\pi/2}$ are identical to the noncommuting phase-difference operators C_{12} and S_{12}, respectively.[159] The classical analog of this is given by the harmonic functions in (3.3.11), namely, $\cos(x_2-x_1)$ and $\sin(x_2-x_1)$ which assume a continuous set of values in the interval $[-1, +1]$. In states with a fixed total number of photons (stationary state), the operators C_{12}, S_{12} have a discrete spectrum.[159] In particular, in the two-mode single-photon state that we are considering, they assume only the two values $\pm 1/2$, i.e., the phase difference has an equal probability of assuming the value $\pi/3$ or $2\pi/3$ for C_{12}, and $\pm\pi/6$ for S_{12}. In view of the connection between C_{12}, S_{12} and the Pauli matrices σ_x, σ_y, we thus have an optical-wave analog of the space quantization of spin 1/2. By increasing to $2j+1$ the number of modes over which the single photon is distributed, we can naturally extend this analogy to the space quantization of an arbitrarily spin j of a single particle.[39]

3.5. Effect of random coincidences

In real experiments, we usually measure the photon coincidence rate,[31] i.e., the rate R_{++} of simultaneous triggering of detectors D_+^a and D_+^b (Fig. 1a), the corresponding rate R_{+-} for D_+^a and D_-^b, and similarly for R_{-+}, R_{--}. If the efficiencies of all four detectors are the same, then symmetry shows that there are only two quantities to be measured, namely, $R_+ \equiv R_{++} = R_{--}$ and $R_- \equiv R_{+-} = R_{-+}$ are identical $R_{+-} = R_{-+}$. The normalized correlator is then

$$E_{\exp} \equiv \langle F \rangle_{\exp} = \frac{R_+ - R_-}{R_+ + R_-}. \tag{3.5.1}$$

Actually, this expression is the same as the 'frequency' definition of the average (3.1.1) because the numerator is multiplied by the time occupied by the measurements is equal to $\Sigma_i F_i$ and the denominator multiplied by the same time is the total number of recorded pairs.

When the detector output pulses due to successively emitted photon pairs begin to overlap, random coincidences begin to appear and are characterized by the rate $R_{\rm acc}$ that is independent of the phase $\varphi = \alpha + \beta$. Consequently, (3.1.3) shows that the recorded rates are given by

$$\begin{aligned} R_+ &= R_{\rm acc} + R' \cos^2(\varphi/2), \\ R_- &= R_{\rm acc} + R' \sin^2(\varphi/2), \end{aligned} \tag{3.5.2}$$

where R' is the maximum rate of 'true' coincidences (in particular, $R' = R_+$ when $R_{\rm acc} = 0$ and $\varphi = 0$).

Substituting (3.5.2.) into (3.5.1), we obtain

$$E_{\exp} = V \cos \varphi, \tag{3.5.3}$$

where

$$V = [1 + (2R_{\rm acc}/R')]^{-1}. \tag{3.5.4}$$

Recalling that the visibility V is not equal to unity, we find that the Bell observable is given by

$$\langle S \rangle_{\exp} = \sqrt{2} V, \tag{3.5.5}$$

so that BI is violated only for $V > 1/\sqrt{2} \approx 0.71$ (Refs. 39, 59, and 160).

We must now elucidate the dependence of V on the parametric conversion efficiency and the quantity $Q \equiv 2\Delta\nu T \sim T/\tau_{\rm coh}$ where $\Delta\nu = \Delta\omega/2\pi$ is the frequency bandwidth of the generated radiation and T is the time constant of the detector. If the radius of the detector aperture is $R_{\rm det} \ll \rho_{\rm coh}$ where $\rho_{\rm coh}$ is the coherence length of the received radiation, the photon counting rate R, i.e., the probability of detection of a photon per unit time, is given by

$$r = \eta' N/T, \quad \eta' = \eta (T/\tau_{\rm coh})(R_{\rm det}/\rho_{\rm coh})^2, \tag{3.5.6}$$

where η is the quantum efficiency of the detector and

$$N = \langle n \rangle = \mathrm{sh}^2\, \tau \tag{3.5.7}$$

is the mean number of photons per mode at the center of the band $\Delta\nu$ (see Appendix III). The relation given by (3.5.6) applies to each of the four detectors.

Since the mean interval of time occupied by the recorded photons is R^{-1}, the probability of a count being recorded in time $T \ll R^{-1}$ is T/R^{-1}, and the rate of random coincidences in the channels is equal to the ratio of the square of this probability and T:

$$R_{\rm acc} = R^2 T. \tag{3.5.8}$$

It can be shown[154] that

$$R'/R_{\rm acc} = \mu^2/f(Q), \tag{3.5.9}$$

where $f(Q) = 2Q^2/[1 - \exp(-2Q) + 2Q]$, i.e., $f(Q) = 1$ for $Q \ll 1$ and $f(q) = Q$ for $Q \gg 1$. The parameter μ is determined by the relative nonstationary correlator (see Appendix III):

$$\mu \equiv \frac{|\langle ab \rangle|}{N} = \coth \tau = \left(\frac{N+1}{N}\right)^{1/2}. \tag{3.5.10}$$

Actually, this is a measure of the nonclassical character of the parametric noise (the nonclassical parameter was introduced in a more general form in, for example, Ref. 161). For classical averaging, it follows from the Cauchy–Schwartz inequality that[50]

$$|\langle ab \rangle|^2 \leqslant \langle a^* a \rangle \langle b^* b \rangle, \tag{3.5.11}$$

so that

$$\mu_{\rm class} \leqslant 1. \tag{3.5.12}$$

The condition $\mu^2<1$ also determines the range of existence of the Glauber–Sudarshan P-distribution[154] that plays the part of the probability distribution for the amplitudes a, b (Ref. 43). Moreover, the parameter μ is simply related to the photon bunching parameter:

$$g\equiv\langle a^+ab^+b\rangle/\langle a^+a\rangle\langle b^+b\rangle=1+\mu^2. \tag{3.5.13}$$

For Poissonian and Gaussian statistics, g is equal to 1 and 2, respectively.

For parametrically scattered radiation, the strong inequality $\mu^2\gg1$ means that the true coincidence rate is much greater than the delayed (random) coincidence rate. This effect was detected experimentally by Burnham and Weinberg[162] in 1970; the existence of the nonstationary correlators $\langle ab\rangle$ and their superclassical magnitude were predicted in Ref. 163.

We note that the radiations emitted by two amplifiers with a common pump are not independent: the correlators $\langle a_1b_1\rangle$ and $\langle a_2b_2\rangle$ have definite phases that are related to the pump phase; according to (3.3.4) this leads to intensity interference.

In accordance with (3.5.9) and (3.5.10),

$$\frac{R'}{R_{\rm acc}}=\frac{N+1}{Nf(Q)}=\frac{\operatorname{cth}^2\tau}{f(Q)}, \tag{3.5.14}$$

and, if we use (3.5.4), we find that the visibility is given by

$$V=[1+(2f(Q)/\mu^2)]^{-1}$$
$$=(N+1)/(2Nf(Q)+N+1). \tag{3.5.15}$$

In a typical experiment with continuous unfocused pump $N\approx\tau^2\sim10^{-8}$ and $Q\sim1000$, so that $V\approx1$. For a stronger pump (pulsed and/or focused), or if we use a resonator, the parametric conversion efficiency is found to be higher and N increases. When $N\gg1$ (parametric superluminescence or generation), we have the following limit even for small $Q\ll1$:

$$V=\frac{N+1}{3N+1}=\frac{1}{1+2\tanh^2\tau}\approx\frac{1}{3}. \tag{3.5.16}$$

This is a typical value of visibility for intensity interference in the case of chaotic sources of light.[33,34]

According to (3.5.5) and (3.5.15), BI is violated for

$$\mu^2>\frac{2f(Q)}{\sqrt{2}-1}\approx4.9f(Q) \tag{3.5.17}$$

or

$$N<\frac{\sqrt{2}-1}{2f(Q)+1-\sqrt{2}}. \tag{3.5.18}$$

In particular, for $Q\ll1$, we have $N<0.26$ (Ref. 39).

If we use (3.5.12) and (3.5.15), we obtain

$$V_{\rm class}\leqslant[1+2f(Q)]^{-1}. \tag{3.5.19}$$

Usually, the values encountered in experiments are $V\simeq1$ and $Q\sim10^3$, so that this limit is exceeded by several orders of magnitude! Consequently, there is a visibility range between (3.5.19) and $1/\sqrt{2}$ in which the classical statistical model that satisfies the Cauchy–Schwartz inequality is contradicted, but BI is still not violated.[19,50,160] The difference between the predictions of classical and quantum theories is thus much clearer in this case than in the Bell theorem. However, the latter is much more fundamental because it does not depend on any particular model of the experiment and is based on very general premises.

We also note that, for two-photon intensity interference based on the Francon scheme,[33,154,156,164,165] Ou and Mandel[166] obtained the following visibility limit from general stochastic theory of ergodic random processes with $T\gg\tau_{coh}$:

$$V\leqslant[(T/\tau_{\rm coh})-1]^{-1}. \tag{3.5.20}$$

They emphasize that the fact that this conflict with experiment can be removed at the cost of abandoning the ergodic hypothesis. The situation resembles the discussion of BI violation[9,93,152–153] in which an alternative to QT was indicated, namely, a negative joint distribution function.[152]

Our model result (3.5.19) satisfies (3.5.20). On the other hand, parallel classical and quantum descriptions of the Francon scheme[154,156] lead to the conclusion that classical visibility obeys an inequality such as (3.5.20) which follows from (3.5.11) in the case of random amplitudes.

3.6. Interference of squeezed noise

Consider a modification of the experiment illustrated in Fig. 1a which can be performed both in the optical (including infrared) and radio-frequency ranges.[33,154] Its distinguishing feature is the presence of 'seed' Gaussian noise at the input of both parametric amplifiers with intensity N_0 photons per mode, i.e., spectral brightness $B_{\omega\Omega}=\hbar cN_0/\lambda^3$. If $N_0\gg1$, the amplifier quantum noise which appears as a result of the vacuum 'seeding' can be neglected because the fact that the operators do not commute ceases to have an effect on the final result. All that remains at the interferometer inputs is the squeezed Gaussian radiation transformed by the phase-sensitive parametric amplifiers, and the classical description becomes admissible. The properties of this type of radiation are analogous to those of squeezed quantum noise: the variances of fluctuations in the two quadratic components are different in the degenerate amplifier regime ($\omega_a=\omega_b$). It can therefore be called 'classical squeezed noise'. We recall that the phrase 'quadrature components' is understood in quantum theory to refer to the operators (Hermitian operators) $X=(a+a^+)/2$, $Y=(a-a^+)/i2$, whereas in classical theory they are the real and imaginary parts of the slowly-varying complex amplitude of the wave.

This analogy was first noted by S. A. Akhmanov *et al.*[167–170] The classical theory of parametric conversion of noise was developed by S. A. Akhmanov *et al.*, Yu. E. D'yakov, and A. S. Chirikin *et al.*[167,171–176]

If the classical condition $N_0>1$ is used only in the final formulas, we can follow the continuous transition from quantum to classical theory. It is striking that the only difference between the results is then the lower visibility in the classical case, which is a consequence of the assumed

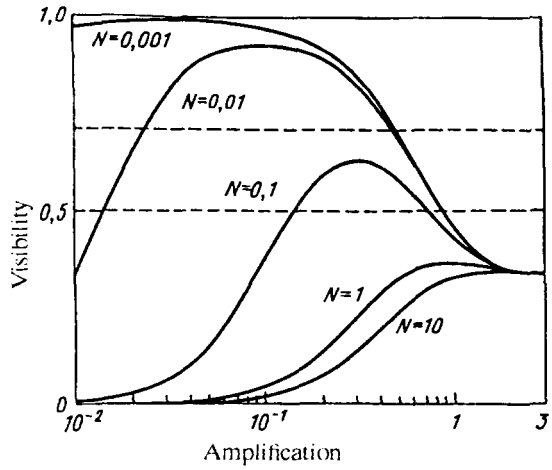

FIG. 4. Visibility V as a function of the growth rate of parametric amplification, τ, of the source of radiation for different intensities N (photons per mode) at the amplifier input. The dashed horizontal lines are the visibility upper limits above which BI$(1/\sqrt{2})$ and the prediction $(1/2)$ of the classical model (3.3.8) are found to fail.

linearity of the conversion process.[33,154] We note that three-photon or multiple-photon parametric amplifiers are nonlinear and occassionally produce qualitative differences between the properties of spontaneous and amplified noise.[35,177]

It is shown in Appendix III [see (III32) and (III33)] that, when seeding noise of equal intensity N_0 is present at the signal and idle mode inputs, the formula for the mean number of photons per mode (3.5.7) is replaced by

$$N+\tfrac{1}{2}=(N_0+\tfrac{1}{2})\operatorname{ch} 2\tau, \tag{3.6.1}$$

and the nonstationary correlator at the output of each of the amplifiers becomes

$$M\equiv\langle ab\rangle=(N_0+\tfrac{1}{2})\sinh 2\tau. \tag{3.6.2}$$

Let us suppose that $Q\sim T/\tau_{\rm coh}\ll 1$, which will enable us to ignore random coincidences between photon pairs due to the fact that we are using 'multimode' detectors. According to (3.5.15) and (3.5.10), the interference visibility is then given by

$$V=\frac{M^2}{2N^2+M^2}=\left\{1+\frac{2}{(2N_0+1)^2}\left[(N_0+1)\tanh\tau+N_0\coth\tau\right]^2\right\}^{-1}. \tag{3.6.3}$$

Figure 4 shows graphs of this function. They readily reveal the conditions under which the Bell and Cauchy–Schwartz inequalities are violated.

The reason for the reduction in visibility with increasing N_0 is the same as in the last Section, namely, random coincidences due to photon-pair overlapping.

In practice, the seeding noise can be accompanied by multimordality of the detectors ($Q>1$) and the visibility must be calculated from (3.5.10), (3.5.15) and (3.6.1) and (3.6.2) without assuming that $T\ll 1$.

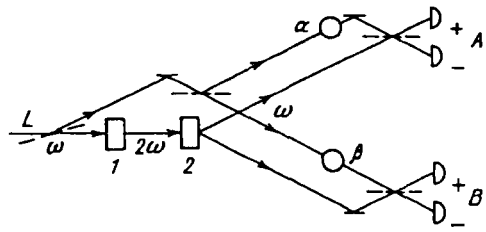

FIG. 5. The principle of the homodyne interferometer: *1*—frequency doubler for the laser L, *2*—parameteric frequency down-converter, operating in the frequency-degenerate regime. The remaining notation is the same as in Fig. 1.

The classical visibility limit as $N_0\to\infty$ follows from (3.6.3) [compare this with (3.5.16)]:

$$V=\frac{1}{1+2\coth^2 2\tau}\leqslant\frac{1}{3}. \tag{3.6.4}$$

Thus, both in the quantum case ($N_0\ll 1$) and in the above classical model ($N_0\gg 1$), the visibility is $V=1/3$ for $\tau\gg 1$, which, by the way, is typical for Gaussian interference as well.[33,34]

3.7. Homodyne detection in intensity interference

In the scheme of Fig. 1, observer A receives two waves a_1 and a_2 with independent random phases x_1 and x_2. The difference x between these phases is measured by the interferometer. Correlation with measurements by observer B will clearly require correlation between the phase differences x and y in the channels, and this is expressed by (3.3.6a). This interpretation is also valid for the quantum description of the process provided the phase-difference operators are suitably defined as indicated in Appendix I.

It is natural in this situation to consider measurement of the 'absolute' values of the phase differences rather than the differences themselves. The former are measured from a particular phase reference, i.e., the homodyne reference wave. This could be exploited to avoid the use of more than one parametric sources of radiation. A possible solution of this kind is illustrated in Fig. 5. It is clear that the homodyne phases in channels A and B must be correlated, so that we again have to have two communication lines to the source for each observer. According to the interferometer classificiations given in Refs. 33 and 34, this is four-mode intensity interference.

Let us now denote the signal and idle mode operators of a given parametric converter by $a\equiv a_1$ and $b\equiv b_1$, and the homodyne mode operator by $c\equiv a_2=b_2$. Analogous schemes were examined in Refs. 54–59 and 178. The adjustable phases α and β can be regarded as additional homodyne phases. We recall that the phase of the nonstationary correlator $\langle ab\rangle$ include the pump phase φ_0 which we usually assume to be zero. Hence, the homodyne phase must be related to φ_0. A single master laser is, of course, used in a real experiment.

Suppose that the homodyne mode is a coherent state with amplitude $\sqrt{2}z$ According (3.4.3) and (3.4.4), we have

$$\langle A_\alpha B_\beta\rangle = |z|^2\langle ab\rangle e^{i\varphi}+\text{h.c.}, \quad \varphi=\alpha+\beta. \tag{3.7.1}$$

where for simplicity the homodyne phases are set to zero.

The normalizing divider K now takes the form [cf. (III36)]

$$\begin{aligned} K^2 &\equiv \langle (n_a+n_c)(n_b+n_c)\rangle \\ &= \langle n_a n_b\rangle + \langle n_a+n_b\rangle\langle n_c\rangle + \langle n_c^2\rangle \\ &= N^2+M^2+2N|z|^2+|z|^4, \end{aligned} \tag{3.7.2}$$

where the functions $N(\tau,N_0)$ and $M(\tau,N_0)$ are defined by (III 32) and (III 33).

We recall that this is the universal normalization in both classical and quantum approaches provided the following conditions are satisfied: each pair photon count is assigned the value $+1$ or -1 in the photon counting regime, depending on which detector records it; in continuous observations, the radiation intensities are held constant at all interfermometer inputs.

With the above normalization, we have

$$E_\varphi = \frac{\langle A_\alpha B_\beta\rangle}{K^2} = V\cos\varphi, \tag{3.7.3}$$

$$V = 2M|z|^2/K^2. \tag{3.7.4}$$

Optimizing the homodyne amplitudes so that

$$|z|^4_{\text{opt}} = N^2+M^2, \tag{3.7.5}$$

we obtain the maximum visibility

$$V_{\max} = M/[N+(N^2+M^2)^{1/2}]. \tag{3.7.6}$$

If we have vacuum at the input of the parametric converter ($N_0=0$) and, there is no 'seed', we have $N=\sinh^2\tau$, $M=(1/2)\sinh 2\tau$, and

$$V_{\max} = [\tanh\tau + (1+\tanh^2\tau)^{1/2}]^{-1}. \tag{3.7.7}$$

This is the result reported by Tan *et al.*[59]

In the case of parametric scattering, $\tau \ll 1$ and we have $|z|^4_{\text{opt}} = N/4 = \tau^2/4$ and the visibility is $V=1$, which is the pure quantum case. If we increase the conversion efficiency to

$$\tau = \tfrac{1}{2}\operatorname{arccosh}\tfrac{9}{7} \approx 0.37,$$

we obtain the HVT limit $V=1/\sqrt{2}$.

Assuming that $N_0 \gg 1$, we obtain the classical squeezed noise for which $N=N_0\cosh 2\tau$ and $M=N_0\sinh 2\tau$, so that

$$\begin{aligned} V_{\max} &= \frac{\sinh 2\tau}{\cosh 2\tau + (\cosh 4\tau)^{1/2}} = \tau, \quad \tau \ll 1, \\ &= 1/(1+\sqrt{2}), \quad \tau \gg 1 \end{aligned} \tag{3.7.8}$$

for

$$|z|^4_{\text{opt}} = \cosh 4\tau. \tag{3.7.9}$$

Thus, effectively squeezed classical noise gives a harmonic interference structure with 41% visibility in the homodyne field. Of course, BI is not then violated.

4. Experiments with three observers

A natural though not very realistic generalization of the above experiment is shown[19,21] in Fig. 1b. The difference as compared with Fig. 1a is that there is now a further two-mode channel. This is an implementation of the GHZ idea[14,15] which provides a demonstration of the violation of the Bell inequality $|\langle S_3\rangle| \leqslant 1$ in a relatively small number of samplings, since S_3 reaches its maximum value of 2 (instead of $\sqrt{2}$ for $\langle S_2\rangle_\psi$) in the case of complete correlation, i.e., for extremal values of the components of its four correlators of the form $P \equiv \langle ABC\rangle = \pm 1$. Statistical analysis of the results of individual (threefold) coincidences is thus unnecessary, at least in principle. This also brings us closer to the EPR program which, in its original form, assumes complete correlation between the measured quantities.[1,15]

Moreover, the addition of further two observables C, C' to the original four A,A',B,B' enables us to formulate very clearly a new type of paradox, namely, the Bell theorem without the GHZ inequalities.[14,15] This is a clear example of a transition from quantity (number of observers) to quality (new type of contradiction).

4.1. Six-mode, three-channel interferometers

Figure 1b shows two parametric sources ($k=1,2$) emitting a triplet of photons ($\nu=a,b,c$) where $\omega_a+\omega_b+\omega_c=\omega_0$. In each channel there are adjustable phase delays α, β, γ. There are three light splitters and six photon sources (or simply detectors) $D^\nu_\pm$ with three green ($+$) and three red ($-$) lamps attached to each them for the sake of clarity.

The sources and detectors are not perfect, but this can be overcome, as before, by using a coincidence scheme which accepts only triple events in which three photons ($i=1,2,\ldots$) produce the simultaneous flashing of three lamps in different channels. The observables A_α, B_β, C_γ are assigned the values ± 1 depending on the color of the lamp that has flashed. The pump power is chosen to be low enough to ensure that neighboring photon triplets do not overlap accidentally during the detector time constant T.

Quantum theory predicts the following results that will undoubtedly be confirmed experimentally after the relevant technical difficulties have been overcome.

For a certain value of the resultant phase $\varphi=\alpha+\beta+\gamma$, which we take as our origin $\varphi=\alpha=\beta=\gamma=0$, we have complete correlation, i.e., each sampling i produces the flashing of an odd number (one or three) of green lamps and an even number (2 or 0) of the red lamps. When $\varphi=\pm\pi$, the reverse picture obtains. Thus,

$$F_{0i}=1, \quad F_\pi=-1, \tag{4.1.1}$$

where

$$F_{\varphi i} \equiv A_{\alpha i}B_{\beta i}C_{\gamma i}. \tag{4.1.2}$$

We recall that for fully correlated measurements in two-observer experiments, no conflict with HVT has been observed. On the other hand, in the above case, the 'experimental' results (4.1.1) are incompatible with classical ideas. This has been clearly demonstrated by Mermin without the use of any mathematics.[118]

We now continue our thought experiment by varying φ and calculating the average multichannel observable F_φ. When the duration of the series of samplings is long enough, the dependence should be harmonic, i.e.,

$$E_\varphi = \langle F_\varphi \rangle_\psi = \cos\varphi \approx \langle F_\varphi \rangle_{\exp} \equiv L^{-1} \sum_{i=1}^{L} F_{\varphi i}. \quad (4.1.3)$$

This is an example of three-photon intensity interference with 100% visibility.

The root mean square deviation of the observable F_φ in QT is

$$\langle \Delta F_\varphi^2 \rangle_\psi^{1/2} = |\sin\varphi|. \quad (4.1.4)$$

By analogy with the two-observer experiment, we use two fixed values of the phases in each channel:

$$\alpha - \alpha' = \beta - \beta' = \gamma - \gamma' = \pi/2 \quad (4.1.5)$$

and select the following combinations:

$$(\alpha',\beta,\gamma), \quad (\alpha,\beta',\gamma), \quad (\alpha,\beta,\gamma'), \quad (\alpha',\beta',\gamma'), \quad (4.1.6)$$

for which we perform four successive series of samplings in which we record

$$\begin{aligned} &F^{(1)} = A'BC, \quad F^{(2)} = AB'C, \\ &F^{(3)} = ABC', \quad F^{(4)} = A'B'C', \end{aligned} \quad (4.1.7)$$

where $A = A_\alpha$, $A' = A_{\alpha'}$, and so on.

According to (4.1.3), the averages are given by

$$\begin{aligned} &E^{(1)} = E^{(2)} = E^{(3)} = \cos(\varphi - \pi/2), \\ &E^{(4)} = \cos(\varphi - 3\pi/2) \end{aligned} \quad (4.1.8)$$

for fluctuations described by (4.1.4) in which $\varphi = \alpha + \beta + \gamma$.

We now impose the additional condition $\varphi = \pi/2$ which is realized, for example, for $\alpha' = \beta' = \gamma' = 0$. This fixes the observables $A = A_{\pi/2}$, $A' = A_0$, and so on (in spin language, these are the Pauli operators $A = \sigma_y^a$, $A' = \sigma_x^a$ (see Appendix I). According to (4.1.4), the observables $F_i^{(m)}$ do not now fluctuate, but assume two possible values, namely, either $+1$ (complete correlation) or -1 (complete anticorrelation).

We now use four separate samplings ($i = 1,2,3,4$) and obtain the twelve values

$$A_1'B_1, \ C_1; \ \ldots; \ A_4', \ B_4', \ C_4'. \quad (4.1.9)$$

Multiplying together the triples in (4.1.2), and using (4.1.8), we obtatin

$$\begin{aligned} &F_1^{(1)} = -1, \quad F_2^{(2)} = -1, \\ &F_3^{(3)} = -1, \quad F_4^{(4)} = 1. \end{aligned} \quad (4.1.10)$$

We note, by the way, that from the standpoint of QT this introduces a particular inconsistency: the right-hand sides contain operators and the left-hand side contain numbers. Strictly speaking, therefore, we have to replace $F^{(m)}$ with $\langle F^{(m)} \rangle_\psi$, and 1 with the unity operator I. Combining the first three inequalities, and subtracting the fourth, we obtain the Bell variable for the three-channel experiment[20]

$$\begin{aligned} S &\equiv \tfrac{1}{2}(F^{(1)} + F^{(2)} + F^{(3)} - F^{(4)}) \\ &= \tfrac{1}{2}(A_1'B_1C_1 + \ldots - A_4'B_4'C_4') \\ &= -2. \end{aligned} \quad (4.1.11)$$

If on the other hand, we follow Mermin[25] and multiply together all four equations in (4.1.10), we obtain the GHZ observable

$$Z \equiv \prod_{m=1}^{4} F^{(m)} = A_1'B_1C_1 \ldots A_4'B_4'C_4'. \quad (4.1.12)$$

These are 'experimental' results. For example, the first eleven quantities in (4.1.9) could be equal to -1 and the twelfth is $C_4' = +1$. We now compare the 'measured' S and Z with HVT predictions.

4.2. Bell inequalities for three observers

The experimental values A_i, A_i', and so on, are predetermined in HVT by means of determined functions $A(\lambda_i)$, $A'(\lambda_i)$,...., where $\lambda_i = \lambda(t_i)$ is a set of variables (whose number may reach, say, 10^{23}) that completely specifies the instantaneous state of the experimental arrangement. In particular, it also determines the time t_i at which the successive event takes place and the twelve specific values of the six observables A, B, C, A', B', C' in the above experiment with four events. In each sampling, a given photon a is thus forced, say, upward in the interferometer channel with phase α, so that we have the event $A_\alpha = +1$. Moreover, the same photon also carries information on how it would behave if we were to encounter another phase α'.

Thus, we obtain, at least in principle, the determined function

$$\begin{aligned} S(\lambda_1,\lambda_2,\lambda_3,\lambda_4) = \tfrac{1}{2}[&F^{(1)}(t_1) + F^{(2)}(t_2) + F^{(3)}(t_3) \\ &- F^{(4)}(t_4)] = \tfrac{1}{2}[A'(\lambda_1)B(\lambda_1)C(\lambda_1) \\ &+ A(\lambda_2)B'(\lambda_2)C(\lambda_2) + A(\lambda_3)B(\lambda_3) \\ &\times C'(\lambda_3) - A'(\lambda_4)B'(\lambda_4)C'(\lambda_4)], \end{aligned} \quad (4.2.1)$$

that describes the outcomes of four samplings with different adjustable phases α, β, γ, where $\lambda_i = \lambda(t_i)$ We note that it does not follow from (4.2.1) that, for example, we must have $C(\lambda_1) = C(\lambda_2)$, since λ_1 need not in general be equal to λ_2. HVT introduces restrictions on only the average experimental data.

As in Sec. 3.2, let us evaluate the average of S over the times t_i:

$$\langle S\rangle_t=\frac{1}{2}(\langle F^{(1)}\rangle_t+\langle F^{(2)}\rangle_t+\langle F^{(3)}\rangle_t-\langle F^{(4)}\rangle_t), \tag{4.2.2}$$

where $\langle ...\rangle_t$ represents averaging over the set of realizations at different instants of time.

We shall suppose that the evolution of $\lambda(t)$ is a random argodic process with a distribution ρ_λ, and evaluate the average over the ensemble of experimental systems, e.g.,

$$\langle F^{(1)}\rangle_t\to\langle F^{(1)}\rangle_\rho\equiv\int A'_\lambda B_\lambda C_\lambda\rho_\lambda \mathrm{d}\lambda. \tag{4.2.3}$$

The result is

$$\langle S\rangle_\rho=\int S_\lambda\rho_\lambda \mathrm{d}\lambda, \tag{4.2.4}$$

where all the λ's are now equal:

$$S_\lambda\equiv\tfrac{1}{2}(A'_\lambda B_\lambda C_\lambda+A_\lambda B'_\lambda C_\lambda+A_\lambda B_\lambda C'_\lambda-A'_\lambda B'_\lambda C'_\lambda). \tag{4.2.5}$$

The observables under the integral sign in (4.2.1) thus lose their time subscripts, so that they can group themselves in an arbitrary manner.

Let us now define two auxiliary variables S_2 and S'_2 that differ by the interchange of the primed and unprimed observables (for the moment, we omit the index λ):

$$S_2\equiv\tfrac{1}{2}[A(B+B')+A'(B-B')],$$

$$S'_2\equiv\tfrac{1}{2}[A'(B'+B)+A(B'-B)]. \tag{4.2.6}$$

It is clear that S_2 and S'_2 can only assume the values ± 1 for dichotomic observables [see (4.2.6)].

We now form the quantity

$$S\equiv S_3\equiv\tfrac{1}{2}[S_2(C+C')+S'_2(C-C')]=\pm 1, \tag{4.2.7}$$

where either $C=C'$ and $S=S_2C=\pm 1$ or $C=-C'$ and $S=-S'_2C'=\pm 1$. This result is valid for any set of phases α, β, γ and α', β', γ', and not only for (4.1.5). The sole essential point is the assumption that (4.1.2) is local, i.e., that A is independent of β, γ, and so on, which enables us to group the observables as in (4.2.6) and (4.2.7).

In classical probability theory, the measure ρ_λ is always non-negative and $\int\rho_\lambda \mathrm{d}\lambda=1$, so that the integral in (4.2.4) cannot lie outside the interval $[-1,+1]$:

$$|\langle S\rangle_\rho|\leqslant\int|S_\lambda|\rho_\lambda \mathrm{d}\lambda=1. \tag{4.2.8}$$

Thus, in HVT,

$$|\langle S\rangle_\rho|=\tfrac{1}{2}|E^{(1)}+E^{(2)}+E^{(3)}-E^{(4)}|\leqslant 1. \tag{4.2.9}$$

This is in fact the Bell inequality for three observers.[20,71] It is also valid in experiments with continuous observables A, A', B, B', C, C' whose numerical values do not exceed unity [see (3.2.9)].

We note that, if we remove the subscripts i from (4.2.1), we obtain (4.2.5) which leads to the erroneous conclusion that the Bell inequality is satisfied in each experimental series of four events. Actually, the unaveraged Bell observable (4.2.1) lies in a wider interval of possible values, namely, $[-2, 2]$.

The BI can also be derived without explicitly involving HVT by starting with the assumption that, immediately before each successive triple event, the future readings of the detectors are predetermined for two variants of the phase delay (in the present case 0 and $\pi/2$) in each of the channels. This is equivalent to an it *a priori* instruction to a photon to move up or down in the beamsplitter for both phase delays (0 and $\pi/2$). A simple radio-frequency model of this situation is described in Sec. 4.4. The 'instruction' is then carried by the random phase difference between the signals from the two parametric generators.

It is clear from the derivation of the BI that, to remove the contradiction between (4.2.9) and the 'experimental value' $\langle S\rangle_{\exp}=\langle S\rangle_\psi=-2$ [cf. (4.1.11)], it is sufficient to abandon one of two assumptions, namely, (a) locality [if $A_\alpha\to A_{\alpha\beta\gamma}$, $B_\beta\to ...$, then (4.2.6) and (4.2.7) are violated] or (b) nonnegativity of ρ_λ [which violates (4.2.8)]). Both alternatives lead to equally unpleasant contradictions, e.g., conflict with relativity theory,[142] but judging by generally accepted terminology, most physicists prefer 'nonlocality'. The mathematical equivalence of these alternatives was demonstrated by Wodkiewicz[153] who considered the example of two spins in a singlet state (see also Refs. 93 and 152).

There is, finally, a third way of escaping from this difficulty: we can abandon the assumed existence of a joint distribution function for the six observables, which can be constructed from ρ_λ. This approach is adopted in the Copenhagen interpretation of quantum mechanics: primed and unprimed operators do not commute, e.g., according to (3.4.4)

$$[A,A']\equiv[A_{\pi/2},A_0]=2i(a_2^+a_2-a_1^+a_1), \tag{4.2.10}$$

so that there is no meaning to the distribution $P(A,A')$.

Other variants of classical theories that allow violation of BI are discussed in the extensive literature now available (see, for example, Refs. 8 and 9). Each approach involves the abandonment of some firmly established physical principle.

As in the case of two-observer experiments, random coincidences limit the measured correlators to the interval $\pm V$, so that, for example, experiment can yield $E^{(1)}=E^{(2)}=E^{(3)}=-V$, $E^{(4)}=+V$. We then have $|\langle S\rangle_{\exp}|=2V$ and the BI is violated only if the visibility V (relative number of values of $F_i^{(m)}$ with the 'wrong' sign) exceeds 50%. We recall, finally, that the corresponding threshold in the two-channel variant was 71%.

4.3. Bell theorem without inequalities

This theorem, or the GHZ paradox,[14,15] can be briefly formulated as follows. According to the definition given by (4.1.12), each local observable appears twice with the same adjustable phase in the composite observable Z:

$$Z_{\exp}=(A'_1A'_4)(A_2A_3)...(C'_3C_4)=-1. \tag{4.3.1}$$

The phases and combinations of them are chosen in accordance with (4.1.5) and (4.1.6), and (4.1.10) is satisfied. The subscripts in (4.3.1) identify the sampling numbers. It is traditional to consider that the GHZ paradox can be demonstrated in four identical samplings, i.e., without any statistical analysis of the results of numerous measurements, which distinguishes it from the 'usual' BI. Actually, if we discard the indices in (4.3.1), we obtain

$$Z=(AA'BB'CC')^2=+1. \qquad (4.3.2)$$

This is in conflict with the right-hand side of (4.3.1) which summarizes the quantum-mechanical analysis.

We shall now try to elucidate the meaning of the above procedure, i.e., the process of ignoring the sampling number in the light of the contradiction that has arisen between the 'experimental' QT prediction (4.3.1) and the HVT prediction given by (4.3.2).

The multichannel composite GHZ variable Z corresponds in HVT to the following single-valued (determined) function of hidden parameters:

$$Z(\lambda_1,\lambda_2,\lambda_3,\lambda_4)$$
$$=A'(\lambda_1)B(\lambda_1)C(\lambda_1)...A'(\lambda_4)B'(\lambda_4)C(\lambda_4) \qquad (4.3.3)$$

where λ_i is the set of values of the hidden parameters at the time of the ith sampling, e.g., $A_i=A(\lambda_i)$ is the specific value of the observable $A=A_\alpha$ in this realization.

The set of values λ_1 predetermines in HVT not only the actually observed variables in a particular sampling, e.g., $A'_1 = B_1 = C_1 = \pm 1$, but also the three unobserved quantities $A_1 = B'_1 = C'_1 = \pm 1$. This is in contrast to QT in which this possibility is excluded by the complementarity principle, so that the symbol $\langle AA'\rangle$ is simply meaningless because the phase delay (or the orientation of the magnet in the spin experiment) cannot simultaneously have two different values α and α'. In other words, HVT postulates a single-valued correspondence between the point λ_1 in continuous phase space $\Lambda\equiv\{\lambda\}$ of the source and the values of the six observables A, A', B, B', C, C', e.g., $\sigma_1=(+++++++)$. It is clear that the same set σ_1 is generated by a subset of points (subspace) in the space Λ, which we shall denote by Λ_1, where $\lambda_1\in\Lambda_1$.

The total number of different sets σ_k for $2N=6$ dichotomic observables is $2^6=64$. The set σ_1, σ_2, ..., σ_{64} forms a discrete set of observables $\{\sigma\}\equiv\Sigma$.

Another point $\sigma_2\neq\sigma_1$ in Σ, e.g., $\sigma_2=(-++++++)$ corresponds to a different subspace Λ_2. It is important to note that Λ_1 and Λ_2 have no common points (they do not intersect), i.e., $\Lambda_1\cap\Lambda_2=\varnothing$ since otherwise the correspondence between Λ and the set Σ will not be single valued: a given cause $\lambda(\lambda\in\Lambda_1\cap\Lambda_2)$ will then generate two mutually exclusive consequences, σ_1 and σ_2, which is in conflict with Laplacian determinism.

Thus, each 'point' σ_k in the set Σ has its own subspace $\Lambda_k=k=1, 2, ..., 64$ in Λ. Some of these subspaces can be empty, e.g., in the next Section we consider a model in which half of the 64 points are not realized (Fig. 6). In the language of set theory, we have a single-valued function f or a mapping of Λ onto Σ which is represented by

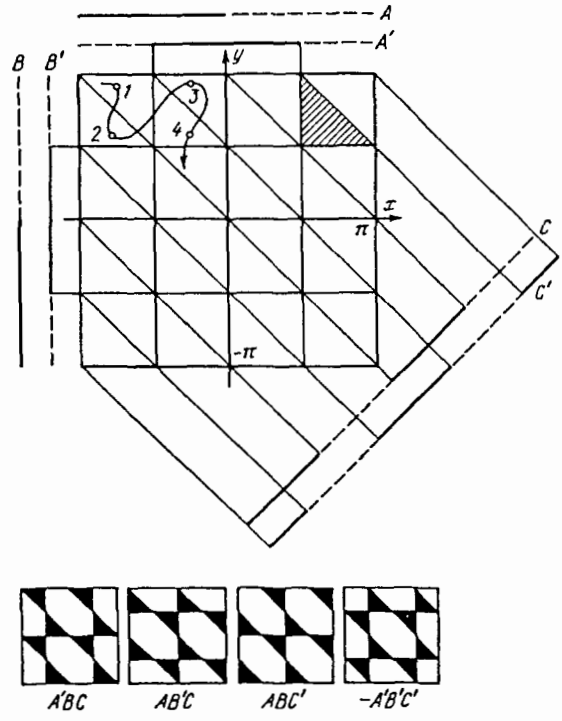

FIG. 6. Diagrams illustrating the mapping of the random-phase space (x,y) occupying the square $[\pi,+\pi]^2$ onto the set of values of the six dichotomic variables A, A', B, B', C, C'. Thick solid lines correspond to $+1$ and dashed lines to -1. The shaded triangle in the phase subspace gives -1 for all the observables. The wavy line is a possible trajectory representing the evolution of phases in time. The black triangles in the lower diagrams give $+1$ for the BI terms $F^{(1)}$, $F^{(2)}$, $F^{(3)}$, and $-F^{(4)}$, given by (4.1.5)–(4.1.7). It is assumed that $\alpha=\beta=\gamma=\pi/2$, $\alpha',\beta'=\gamma'=0$.

$\{f:\Lambda\to\Sigma\}$; in particular, $\sigma_1=f(\lambda_1)$. The reverse transformation is not single-valued and is represented by $\Lambda_1=f^{-1}(\sigma_1)$; Λ_1 is called the complete prototype of σ_1. The union of subspaces $\cup_k\Lambda_k=\Lambda$ exhausts the entire set of possible values of the hidden variables λ.

In the course of time, the mapping point $\lambda(t)$ in the phase space of the source leaves Λ_k for another region, so that if at the time t_i it crosses Λ_k (if we take into account the time taken by the particle between the source and the detectors), the outcome of the experiment is σ_k.

Suppose that in all four samplings the mapping point $\lambda(t)$ is found in the same subspace Λ_k:

$$\lambda_i\in\Lambda_k, \quad i=1,\ 2,\ 3,\ 4, \qquad (4.3.4)$$

in which case the subscripts in (4.3.3) can indeed be discarded or, more correctly, we can write

$$Z_\lambda=(A_\lambda A'_\lambda B_\lambda B'_\lambda C_\lambda C'_\lambda)^2=+1. \qquad (4.3.5)$$

This possibility cannot be excluded theoretically. In this sense, it may be said that the GHZ paradox does not require repeated samplings. However, in reality, during the time occupied by the individual measurements separated by macroscopic intervals Δt (in optical language $\Delta t\gg\tau_{\rm coh}$, the mapping point is most likely to leave Λ_k. For example, suppose that the region Λ_1 defined above corresponds to

time t_2 and λ_2 corresponds to t_3. In that case, in (4.3.1) we have $A_2=+1$, $A_3=-1$, and (4.3.4)–(4.3.5) are not satisfied. Consequently, four samplings will not reproduce the GHZ paradox, and indeed the measured value $Z=-1$ can always be explained by the violation of (4.3.4), i.e., the necessary condition for discarding the subscripts in (4.3.3).

In general, it is impossible to monitor the validity of (4.3.4). The only way of escaping from this situation is therefore by carrying out a large number of samplings that will indicate whether λ_i enters the same subspace Λ_k for all $i=1, 2, 3, 4$ with a particular probability that exceeds the probability predicted by HVT.

Let us now examine the experimental procedure in greater detail. In four series of samplings, we measure the multichannel observables $F^{(m)}$, $m=1$–4 corresponding to the phase combinations (4.1.5)–(4.1.6). In the ideal situation, the quantum mechanical treatment predicts the complete correlation (4.1.10). However, random coincidences may give $[1-V(L)]$ with 'incorrect' signs. Suppose that for each m the total number of 'correct' answers is much greater than the product of 64 by the number of 'wrong' realizations. We may then reliably conclude that at least four of the 'correct' (one for each series m) had the same subspace Λ_k as the prototype. Consequently, for these four samplings, (4.3.4) is satisfied and the sampling number i in (4.3.1) can be excluded, i.e., (4.3.2) and (4.3.5) can be justified.

Thus, a file of say of $4\times 64\times 10$ realizations should clearly demonstrate the GHZ paradox and thus the inadequacy of HVT. Although this does not require formal averaging, the procedure is nevertheless statistical in character, i.e., it enables us to draw a final conclusion with finite probability. The latter can be increased by increasing the number of realizations and is limited only by random coincidences.

Apart from its exceptional clarity, the GHZ paradox has the further positive property that none of the HVT 'rescalings' proposed by Santos[44,45] to explain experimental violations of BI appears to resolve this paradox.

How then do we escape from the contradiction $-1=+1$? The most obvious solution is (this is the 'minimal' interpretation[10]) the abandonment Laplacian determinism [single-valuedness of $\sigma=f(\lambda)$] and reconciliation with the principle of complementarity in view of the fact that it is impossible to assign *a priori* values to noncommuting observables. This is the standpoint adopted by supporters of the orthodox of Copenhagen interpretation of QT. However, a continuing search has been in progress for different loopholes in the above derivation as a means of justifying new variants of 'objective realism' (see, for example, Refs. 44 and 45).

We shall show later that the QT result $Z=-1$ follows from the operator identity $(AA')^2=-I$ which is possible only for noncommuting operators A and A', i.e., they cannot be measured simultaneously in a single sampling. Moreover, the operator AA' is non-Hermitian, $(AA')^+=A'A\neq A'$, which means that it cannot describe an observed variable.

4.4. Stochastic models of three-channel interference

In complete analogy with Sec. 3.3, we shall now describe the interferometer input field by six complex amplitudes a_k, b_k, c_k. The intensities of all six modes will be assumed to be constant and equal, $n_\nu=1$, and the phases $x_k(t)$, $y_k(t)$, $z_k(t)$ will be subjected to the 'parametric' conditions

$$x_k+y_k+z_k=\text{const}, \quad k=1,2. \tag{4.4.1}$$

The phase fluctuations are transformed by the interferometer into intensity fluctuations [see (3.3.2)]:

$$\begin{aligned} n^a_\pm(t)&=1\pm\cos[\alpha+x(t)],\\ n^b_\pm(t)&=1\pm\cos[\beta+y(t)],\\ n^c_\pm(t)&=1\pm\cos[\gamma-x(t)-y(t)], \end{aligned} \tag{4.4.2}$$

where α, β, γ are adjustable phases and $x(t)=x_2-x_1, y(t)=y_2-y_1$.

The correlator of the photocurrents from the three detectors, e.g., $\langle n^a_+, n^b_+ n^c_+\rangle$ exhibits 'third order' interference with phase $\varphi=\alpha+\beta+\gamma$ and visibility 1/4 [see (3.3.7) and (3.3.9)]. This concludes our discussion of the model with continuous variables, and we now turn to the discrete dichotomic A, B, C.

We shall record the signs of the difference between currents $\Delta n_\nu=n^\nu_+ n^\nu_-$ in the three channels at successive times t_i separated by intervals $\Delta t\gg\tau_{\text{coh}}$. This gives the dichotomic sequences

$$\begin{aligned} A_i&=\text{sign}\cos(\alpha+x_i),\\ B_i&=\text{sign}\cos(\beta+y_i),\\ C_i&\ \text{sign}\cos(\gamma-x_i-y_i). \end{aligned} \tag{4.4.3}$$

The two random phase differences $x(t)$, $y(t)$ or x_i, y_i now play the role of the hidden parameters. Assuming that they are ergodic random functions with a uniform distribution $\rho(x,y)=1/4\pi^2$ in the square $\Lambda=[-\pi,+\pi]^2$, we obtain the following expression for the correlation between the signs of the multichannel observable F_i (see Appendix IV for further details):

$$\begin{aligned} E_\varphi&=\langle F_\varphi\rangle_\rho\\ &=\frac{1}{4\pi^2}\iint_{-\pi}^{\pi}\text{sign}[\cos x\cos y\cdot\cos(x+y-\varphi)]\,dx\,dy\\ &=(1-\tilde{\varphi}^2)/2 \quad \text{for } |\tilde{\varphi}|\leqslant 1,\\ &=[(|\tilde{\varphi}|-2)^2-1]/2 \quad \text{for } 1<|\tilde{\varphi}|\leqslant 2, \end{aligned} \tag{4.4.4}$$

where $\tilde{\varphi}=2\varphi/\pi$. The graph of this function consists of four identical segments or parabolas turned in opposite directions relative to one another. It is shown in Fig. 3b.

The experimental values of E_φ are $\pm 1/2$ for $\varphi=0$ and $\pm\pi$, i.e., complete correlation is not observed in this case [compare this with (3.3.18)].

For the combinations of phases given by (4.1.5) and (4.1.6), the BI is satisfied in the limit:

$$\langle S\rangle_\rho=\tfrac{1}{2}(-\tfrac{1}{2}-\tfrac{1}{2}-\tfrac{1}{2}-\tfrac{1}{2})=-1. \tag{4.4.5}$$

The space Λ of hidden parameters is readily realized in this model and is, clearly, connected with the set Σ of dichotomic variables A, A', B, B', C, C', ($A\equiv A_{\pi/2}$, $A'\equiv A_0$, and so on, in accordance with (4.1.5) which was analyzed in a general form in the last Section. The mapping $\Lambda\to\Sigma$ is illustrated in Fig. 6. The two-dimensional space $\Lambda=\{x,y\}$ occupies the square $[-\pi,\pi]^2$ in the xOy plane. According to (4.4.3)

$$\begin{aligned} A&=-\operatorname{sign}\sin x,\\ B&=-\operatorname{sign}\sin y,\\ C&=\operatorname{sign}\sin(x+y),\\ A'&=\operatorname{sign}\cos x,\\ B'&=\operatorname{sign}\cos y,\\ C'&=\operatorname{sign}\cos(x+y) \end{aligned} \tag{4.4.6}$$

and the four multichannel observables $F^{(n)}$ used to form the Bell and GHZ observables are given by

$$\begin{aligned} F^{(1)}&=-\operatorname{sign}[\cos x\cdot\sin y\cdot\sin(x+y)],\\ F^{(2)}&=-\operatorname{sign}[\sin x\cdot\cos y\cdot\sin(x+y)],\\ F^{(3)}&=\operatorname{sign}[\sin x\cdot\sin y\cdot\cos(x+y)],\\ F^{(4)}&=\operatorname{sign}[\cos x\cdot\cos y\cdot\cos(x+y)]. \end{aligned} \tag{4.4.7}$$

We note that, for given x,y (the same for all $m=1$–4), we have

$$Z\equiv F^{(1)}F^{(2)}F^{(3)}F^{(4)}=+1, \tag{4.4.8}$$

i.e., we have pairs such as $F^{(1)}$, $F^{(3)}$ and $F^{(2)}=F^{(4)}$ and other combinations of m.

According to (4.4.6), half of the $2^6=64$ possible variants of the six observables σ_k have as their prototypes empty subspaces Λ_k, e.g., $\sigma_1=(+++++++)$ are never realized in this model and $\Lambda_1=\varnothing$. The remaining 32 subspaces are indicated in Fig. 6 by the triangles, and the corresponding signs of the observables by straight segments.

The random phase functions $x(t)$ and $y(t)$ enter each of the 32 nonempty subspaces Λ_k an infinite number of times for $t\to\infty$, and uniformly fill the square $\Lambda=[-\pi,+\pi]^2$. It is therefore a relatively simple matter to use Fig. 6 to calculate the correlators $E^{(m)}=\langle F^{(m)}\rangle$ as ratios of sums of areas and subspaces Λ_+ and Λ_-, giving $F=+1$ and $F=-1$, to the total area Λ:

$$\begin{aligned} E^{(1)}=E^{(2)}=E^{(3)}=E_\pi&=\frac{8-24}{8+24}=-\frac{1}{2},\\ E^{(4)}=E_0&=\frac{24-8}{24+8}=\frac{1}{2}. \end{aligned} \tag{4.4.9}$$

In these expressions, the unit measure is the area of the elementary triangle Λ_k in Fig. 6. The resulting numbers are special values or E_φ of the form of (4.4.4), which give (4.4.5), i.e., the 'classical' value of the Bell observable: $\langle S\rangle_\rho=-1$.

The question now is whether it is possible, in the classical model, to reach complete correlation (or anticorrelation) between the observables F for $N\geqslant 3$ or whether this is the exclusive prerogative of the quantum-mechanical treatment. To find the answer to this very legitimate question, consider the nonuniform phase distribution $\rho(x,y)$. Suppose, for example, that the phases of two parametric generators, whilst remaining independent and random, are confined to the shaded triangle in Fig. 6. All six observables A, A',..., C' then assume the single value -1 and, naturally, we have complete anticorrelation, e.g., $F^{(1)}=-1$. However, this model is not consistent with our hypothetical experiment in which A is equal to ± 1 and so on with equal probability.

If, on the other hand, we abandon the independence of the phases x and y, we obtain pair correlation between the channels. Thus, assuming that $x=y$, we obtain $A=B$ and $A'=B'$, which is again in conflict with 'experiment'.

The absence of pair correlation is assured by the factorization of the distribution function:

$$\rho(x,y)=\rho_x(x)\rho_y(y), \tag{4.4.10}$$

and the uniformity of the positive and negative outcomes for each of the six observables, i.e., the symmetry of the functions $\rho_x(x)$ and $\rho_y(y)$ with respect to the points at which the observables change sign.

The following model satisfies these conditions. Suppose the phase differences x and y randomly assume values within the intervals $0\pm\delta\varphi$ and $\pi\pm\delta\varphi$ with $\delta\varphi\ll\pi/2$. This situation occurs when a degenerate parametric generator is repeatedly turned on. It was called phase quantization by Akhmanov *et al.*[171,172] We may therefore suppose that

$$\begin{aligned} \rho(x,y)=\tfrac{1}{4}&[\delta(x)+\delta(x-\pi+0)]\\ &\times[\delta(y)+\delta(y-\pi+0)]. \end{aligned} \tag{4.4.11}$$

The shift of the δ-functions by the infinitesimal positive amount will be convenient when we evaluate the integrals in the course of averaging.

According to (4.4.11), all three phase differences x, y, and $z=-x-y$ display in pairs random and independent jumps by π between realizations. Substituting (4.4.11) in an integral such as (4.4.4), we obtain

$$\begin{aligned} E&=\iint\limits_{-\pi}^{\pi}F(x,y)\rho(x,y)\,dx\,dy\\ &=\tfrac{1}{4}[F(0,0)+F(\pi,0)+F(0,\pi)+F(\pi,\pi)]\\ &=\operatorname{sign}(\cos\alpha\cdot\cos\beta\cdot\cos\gamma)=\pm 1, \end{aligned} \tag{4.4.12}$$

where we now have

$$\begin{aligned} F(x,y)=\operatorname{sign}[&\cos(\alpha+x)\cdot\cos(\beta+y)\\ &\times\cos(\gamma-x-y)]. \end{aligned} \tag{4.4.13}$$

For any α, β, $\gamma, \pm\pi/2$, we thus have complete correlation (or anticorrelation) whose sign can be altered by any of the three observers, e.g., by altering α for fixed β and γ. Of course, BI is satisfied: the phase combinations given by (4.1.5) and (4.1.6) with the addition of the infinitesimal shift $+0$ give $E^{(m)}=+1$ for all $m=1$–4.

Complete correlation (anticorrelation) is thus not specific to the quantum model. Moreover, our classical model gives the same low-order moments as the quantum model:

$$\langle A\rangle=\langle B\rangle=\langle C\rangle=\langle AB\rangle=\langle AC\rangle=\langle BC\rangle=0. \tag{4.4.14}$$

The only observed difference is the Π-shaped dependence of E on each of phase α, β, γ (see Fig. 3c) instead of the harmonic interference curve $\cos(\alpha+\beta+\gamma)$ in QT. These results can be generalized to an arbitrary number of recording channels N. This is discussed in Sec. 5.4.

We also note that the GHZ observable in discrete classical models (with dichotomic observables) is independent of the form of $\rho(x,y)$, so that (4.4.8) is always satisfied. We emphasize once again the conflict between this result and QT in which, in the absence of random coincidences, $Z=+1$ is *never* found to occur [conditions (4.1.5)-(4.1.7) must, of course, be satisfied]. Nevertheless, as was shown in the preceding section, experimental confirmations of this fact will be statistical in character. We shall illustrate this with the help of Fig. 6 by considering the following clear example. Without restricting the distribution function $\rho(x,y)$, we assume that a series of four samplings was used to find the random phases $(x,y)\equiv\lambda$ at times $t_i,=1, 2, 3, 4$ at the points indicated by these numbers in Fig. 6. At points 1 and 2, we record the readings $(A'BC)_1=(--+)$ and $(AB'C)_2=(+--)$, so that $C_1\neq C_2$ and the conditions for the appearance of the GHZ paradox are not met (the four sets $\lambda^{(m)}$ do not belong to one triangle). Actually, according to Fig. 6, $F^{(1)}=F^{(2)}=F^{(3)}=+1$, $F^{(4)}=-1$, so that $Z=-1$ and $S=+2$. Hence, in accordance with HVT, individual series of samplings can give results lying outside the classical frame and requiring that the hidden parameters be identical.

4.5. Quantum theory of three-photon interference

Suppose (see Fig. 1b) that we have three photons a, b, c at the interferometer input, each of which is distributed between two modes $k=1$, 2:

$$|\psi\rangle=\frac{1}{\sqrt{2}}\sum_{k=1}^{2}a_k^+b_k^+c_k^+|0\rangle$$
$$\equiv\frac{1}{\sqrt{2}}(|10\rangle_a|10\rangle_b|10\rangle_c+|01\rangle_a|01\rangle_b|01\rangle_c). \tag{4.5.1}$$

The observables A_α, B_β, C_γ are given by (3.4.4) and the multi-channel operator F_φ is given by

$$F_\varphi\equiv A_\alpha B_\beta C_\gamma=\sigma_-^a\sigma_-^b\sigma_-^c e^{i\varphi}+\text{h.c.}, \tag{4.5.2}$$

where $\sigma_-^a=a_1a_2^+$ and so on, $\varphi=\alpha+\beta+\gamma$, and h.c. represents Hermitian conjugates. Terms that simultaneously include σ_- and σ_+ are omitted.

We now apply the observable operators (3.4.4) to the state vector (4.5.1) and obtain

$$A_\alpha|\psi\rangle=\frac{1}{\sqrt{2}}(e^{i\alpha}a_2^+b_1^+c_1^++e^{-i\alpha}a_1^+b_2^+c_2^+)|0\rangle \tag{4.5.3}$$

and similarly for B_β and C_γ.

Next, we find that the average and variance of the multichannel variable are given by

$$E_\varphi\equiv\langle F_\varphi\rangle=\cos\varphi,\quad \langle\Delta F_\varphi^2\rangle=\sin^2\varphi, \tag{4.5.4}$$

i.e., we have complete agreement with the results obtained for the two-photon interferometer. These results will later be generalized to an arbitrary number N of observers [see (5.3.9)].

The joint distribution of the observables A, B, C differs from (3.1.3) by only a numerical factor due to the increase in the number of observation channels:

$$P_{ABC}^{\xi\eta\zeta}=\langle n_a^\xi n_b^\eta n_c^\zeta\rangle=\tfrac{1}{4}\cos^2\frac{\varphi}{2}\quad\text{for } \xi\eta\zeta=+1,$$
$$=\tfrac{1}{4}\sin^2\frac{\varphi}{2}\quad\text{for } \xi\eta\zeta=-1, \tag{4.5.5}$$

where ξ, η, $\zeta=\pm1$.

We now turn to the justification of the GHZ theorem and the 'experimental' result given by (4.3.1). For the phase combinations given by (4.1.5)–(4.1.6) and chosen for the four modifications of the experiment, we obtain the following expression for the GHZ observable if we use (4.1.7):

$$Z\equiv F^{(1)}F^{(2)}F^{(3)}F^{(4)}=A'BC....C'=-I. \tag{4.5.6}$$

On the other hand, since different operators commute, i.e., $[A,B]=[A',B]=...=0$, we have

$$Z=BB'BB'=(BB')^2=I+B[B',B]B'. \tag{4.5.7}$$

This means that, it is precisely the neglect of the fact primed and unprimed operators (in this case, B and B') do not commute that leads to $Z=+I$, i.e., the HVT conclusion. The GHZ, Bell, and KS theorems all formally follow from the noncommunitivity of the algebra of observable operators.

It is shown in Appendix I [see (I1 9)] that, for the corresponding phases, the operators A and A' or B and B' are practically identical with the Pauli operators σ_y and σ_x. Hence, the essence of the GHZ paradox can be summarized as follows:

$$Z=(\sigma_y\sigma_x)^2=(-i\sigma_z)^2=-I\neq\sigma_x^2\sigma_y^2=I. \tag{4.5.8}$$

We also note that, since Z is a product of a noncommuting operators, the corresponding variable cannot be measured in a single sampling. This in turn means that, in experiments with incomplete correlation, we can compare experimental results with the product of averages

$\langle F^{(1)}\rangle\langle F^{(2)}\rangle\langle F^{(3)}\rangle\langle F^{(4)}\rangle$, but not with the average of the product $\langle F^{(1)}F^{(2)}F^{(3)}F^{(4)}\rangle$, since it is simply impossible to measure the latter.

Next, we turn to a generalization of the above model for the case of amplification by a parametric converter of initial Gaussian noise of intensity $N_0=\langle a_0^+a_0\rangle_0 =\langle b_0^+b_0\rangle_0=\langle c_0^+c_0\rangle_0$. We have already performed the analogous operation for the two-channel variant of the interferometer (see Sect. 3.6). In the limit as $N_0\to\infty$ we have generation by the source of the classical analog of three-photon squeezed light.[35,177]

In the Heisenberg representation, three-mode parametric conversion is described by the equations of motion

$$\begin{aligned} &\mathrm{d}a\mathrm{d}\tau=b^+c^+,\\ &\frac{\mathrm{d}b}{\mathrm{d}\tau}=a^+c^+,\\ &\frac{\mathrm{d}c}{\mathrm{d}\tau}=a^+b^+, \end{aligned} \tag{4.5.9}$$

that are the analogs of (III 2). The nonlinearity of (4.5.9) leads to a number of interesting properties (see, for example, Refs. 33 and 179 and the literature cited therein).

We shall take the solution of (4.5.9) in the form of a perturbation series in τ:

$$a=a_0+\tau b_0^+c_0^++\frac{\tau^2}{2}a_0(1+b_0^+b_0+c_0^+c_0)+\dots. \tag{4.5.10}$$

Analogous expressions can be obtained for b and c.

For an initial vaccuum state ($N_0=0$), the first two orders in τ give the previous results (4.5.4) and (4.5.5) after the appropriate normalization. The third-order corrections enable us to take into account random coincidences whose rate in the three-photon experiment is R^3T^2 [see (3.5.8)] where R is the photon count rate in each channel and T is the 'window' of the coincidence system, largely determined by the detector time constant. From now on, we shall confine our attention to the second-order approximation in τ.

The 'appropriate' normalization of the observable mentioned above is analogous to one that we already know from (3.3.14), (3.5.1), (3.7.2) and (III35):

$$\begin{aligned} &A_\alpha=\frac{n_+^a-n_-^a}{K},\\ &B_\beta=\frac{n_+^b-n_-^b}{K},\\ &C_\gamma=\frac{n_+^c-n_-^c}{K}, \end{aligned} \tag{4.5.11}$$

in which

$$K^3=\langle(n_+^a+n_-^a)(n_+^b+b_-^b)(n_+^c+n_-^c)\rangle. \tag{4.5.12}$$

We recall that the normalization is universal for both classical and quantum models, e.g., averaging over the state (4.5.1) gives $K=1$.

We now substitute (4.5.10) in (4.5.11) and form the normalized correlator

$$E_\varphi\equiv\langle F_\varphi\rangle=V\cos\varphi, \tag{4.5.13}$$

where the interference visibility is given by

$$V\approx[1+(4N_0^3/\tau^2)]^{-1}. \tag{4.5.14}$$

The following intermediate results were used in deriving these relations:

$$\begin{aligned} &\langle(n_+^a-n_-^a)(n_+^b-n_-^b)(n_+^c-n_-^c)\rangle_0\\ &\quad=2\tau^2(1+6N_0+15N_0^2+18N_0^3+9N_0^4)\cos\varphi, \end{aligned} \tag{4.5.15}$$

$$\begin{aligned} &K^3\equiv\langle n_an_bn_c\rangle_0\\ &\quad\approx\langle n_an_bn_c\rangle_0|_{\tau=0}+\langle n_an_bn_c\rangle_0|_{N_0=0}\\ &\quad=2(\tau^2+4N_0^3). \end{aligned} \tag{4.5.16}$$

Since we are using the Heisenberg representation, the average $\langle\dots\rangle_0$ is evaluated over the initial chaotic mixed state characterized by the correlators given by (III 27). The approximation defined by (4.5.14) and (4.5.16) is valid for $N_0\ll1$ and $\tau^2\ll1$, and enables us to avoid exceedingly laborious exact calculations. Mixed terms such as τ^2N_0, $\tau^2N_0^2$, and so on, are neglected in (4.5.16) because, from now on, we shall be interested in values $N_0^3\sim\tau^2\gg\tau^2N_0$, $\tau^2N_0^2,\dots$. The ratio of (4.5.15) and (4.5.16) gives (4.5.13) and (4.5.14) for $N_0\ll1$.

We note that (4.5.13) does not, by far, exhaust the manifold of three or more photon interference of different order,[33–35,157,177] which makes itself felt despite the linearity of the interferometers. For example, in some arrangements, the interference maximum can evolve into a minimum as N_0 increases from zero to infinity.[35,177]

However, let us return to our case. As noted at the end of Sec. 4.2, the upper limit of the visibility V for which BI is still satisfied is 1/2. Hence, according (4.5.14), BI is violated for

$$N_0<(\tau^2/4)^{1/3}. \tag{4.5.17}$$

The fall in visibility with increasing N_0 is again due to random coincidences, but for $N_0\neq0$ they occur even in the interferometer with perfect detectors and coincidence window $T=0$.

We note one further interesting point. If the channel-independent Gaussian noise is fed not to the parametric converter, but directly to the interferometer inputs, there is no interference and the visibility is $V=0$. It is readily verified for the superposition of the quantum state (4.5.1) and the chaotic mixed state with $\langle n_k^\nu\rangle_0=1/2$, $\nu=a, b, c$; $k=1, 2$ (i.e., $\langle n_\nu\rangle_0=1$), the visibility is $V=1/2$ and BI is satisfied. Hence, the addition to each photon in a correlated triple of only one chaotic photon ensures that the pure quantum effect vanishes and BI is violated. This agrees with the analogous conclusion derived in Ref. 161 in a more general case.

As in the case of the two-observer experiments, we can have the homodyne modification of the three-photon experiment, namely, the replacement of one of two paramet-

ric sources of 'triphotons' with a coherent homodyne that provides a general phase reference for the three channels. An obvious transformation of the scheme of Fig. 5 is then performed in accordance with the rule $\to 1b$ (Fig. 1a).

If the conditions introduced in the derivation of (4.5.13) and (4.5.14) are met in the homodyne variant in the experiment, then (4.5.13) remains valid and the visibility is given by

$$V \approx \frac{2\tau|z|^3}{\tau^2+(N_0+|z|^2)^3}, \tag{4.5.18}$$

where $|z|^2 \ll 1$ is the intensity of the homodyne radiation introduced into each channel (assumed to be the same in all channels).

If we optimize the efficiency of the parametric process so that

$$\tau_{\rm opt}=(N_0+|z|^2)^{3/2}, \tag{4.5.19}$$

we obtain the maximum visibility

$$V_{\max}=[1+(N_0/|z|^2)]^{-3/2}. \tag{4.5.20}$$

BI is thus violated for $N_0/|z|^2<0.59$, which corresponds to a reduction in visibility from 1 to 1/2 as N_0 increases from zero.

5. BELL THEOREM FOR *N* OBSERVERS

There is a considerable number of publications[14–20,90,116] on the generalization of the Bell theorem and inequalities to the case of an arbitrary number of correlated particles (usually with spin 1/2). An interesting version of this problem was recently solved by Mermin[16] and in a somewhat improved form by Roy and Singh.[17] Subsequent studies were performed by Ardehali[18] and others.[19,21] In all cases, other than those of Refs. 19 and 21, the treatment was based on the N-particle spin model. The principal result of these investigations is the discovery of the possibility of a nontrivial exponential increase with increasing N in the relative discrepancies between the QT and HVT predictions for the N-particle observable S_N:

$$|\langle S_N\rangle_\psi/\langle S_N\rangle_\rho| \geqslant 2^{(N-1)/2}. \tag{5.1}$$

When $N=2$ and 3, this expression gives the familiar results, i.e., (3.1.12), (3.1.13), (4.1.11), and (4.2.9). However, as $N\to\infty$, Mermin[16] shows that a new quantum effect can appear. It is important to remember, however, that the detection of this effect would require the detection of each of the N particles separately.

We now present a derivation and a discussion of the Bell theorem that leads to (5.1) in the case of an optical interference experiment (a hypothetical experiment as yet).[19,21]

5.1. *N*-channel interferometer

Let us now return to Fig. 1 and replace the two or three observers with an arbitrary number of N-observers. We now use two N-frequency parametric converters that successively, at random times t_i, generate ensembles of N photons, each of which belongs to two modes. The dichotomic observables $A_n=(\alpha_n)=\pm 1$, $m=1, 2,..., N$ are recorded. Let α_n be the phase delay in the n-th channel. In the spin variant of the experiment, the phase delay α_n defines the orientation of the Stern–Gerlach analyzer by analogy with the model described in Appendix II. A homodyne scheme generalizing Fig. 5 is also possible.

Now consider the N-channel observable

$$F(\{\alpha\})\equiv A_1(\alpha_1,t_i)A_2(\alpha_2,t_i)...A_n(\alpha_N,t_i). \tag{5.1.1}$$

QT predicts a harmonic interference modulation with 100% visibility:

$$E_N(\{\alpha\})=\langle F_\varphi\rangle_\psi=\cos\varphi, \quad \varphi=\sum_{n=1}^{N}\alpha_n. \tag{5.1.2}$$

The relations given by (3.1.4) and (4.1.3) have thus received a natural generalization.

For complete correlation ($E_n=1$), it is convenient to establish reference values of the adjustable phases in each of the phases: $\alpha_1=\alpha_2=...=\alpha_N=0$.

Although the experiment becomes less graphic as N increases, the contradiction (5.1) with HVT that it introduces becomes more dramatic.

5.2. Bell inequality for *N* observers

In HVT, the results of samplings are predetermined by determined functions of hidden variables: $A_n(\alpha_n,\lambda_i)=\pm 1$, $\lambda_i=\lambda(t_i)$. To simplify our notation, we shall not in future explicitly indicate this dependence on λ.

Meeasurements in each channel will be performed for two values of the phases, α_n and α'_n, and we shall place two observables in correspondence with these phases, namely, $A_n\equiv A_n(\alpha_n)$ and $A'_n\equiv A_n(\alpha'_n)$, which also assume the values ± 1.

The next to arise is the problem of finding the N-channel Bell observables S_N that would produce the maximum conflict between QT and HVT, as for S_2 and S_3. The obvious way of solving this problem is to continue the iteration procedure, which will enable us to find S_3 in the form of (4.2.7):

$$\begin{gathered} S_N=\tfrac{1}{2}S_{N-1}(A_N\pm A'_N)\pm\tfrac{1}{2}S'_{N-1}(A_N\mp A'_N), \\ S_0=S'_0=1, \quad N=1,\ 2,\ldots; \end{gathered} \tag{5.2.1}$$

where S'_N differs from S_N by the interchange of the primes: $A_n\leftrightarrow A'_n$. We have used a more general configuration of signs as compared with (3.2.6) and (4.2.7), but this will not upset the BI established below. The only important point is that the signs inside the parenthesis should always be opposite. In particular, we shall need the following combinations:

$$(-++),\quad (++-),\quad (--+),\quad (+--) \tag{5.2.2}$$

for $N=1$, 2, 3, 4 with period 4 (mod 4), respectively. This provides us with a compact form of S_N as in (5.2.13) and (5.2.14). For lower N, we have

$$S_1=A_1, \quad S'_1=A'_1, \tag{5.2.3}$$

$$S_2=[S_1(A_2+A_2')+(S_1'(A_2-A_2')]/2$$
$$=(A_1A_2+A_1'A_2+A_1A_2'-A_1'A_2')/2$$
$$\equiv(A^2+2AA'-A'^2)/2, \quad (5.2.4)$$
$$S_3=[S_2(A_3-A_3')-S_2'(A_3+A_3')]/2$$
$$=(A_1A_2A_3-A_1A_2'A_3'-A_1'A_2A_3'-A_1'A_2'A_3)/2$$
$$\equiv(A^3-3AA'^2)/2, \quad (5.2.5)$$
$$S_4=[S_3(A_4+A_4')-S_3'(A_4-A_4')]/2$$
$$=(A^4+4A^3A'-6A^2A'^2-4AA'^3+A'^4)/4, \quad (5.2.6)$$
$$S_5=[S_4(A_5-A_5')+S_4'(A_5+A_5')]/2$$
$$=(A^5-10A^3A'^2+5AA'^4)/4, \quad (5.2.7)$$

where $KA^nA'^{N-n}$ represents symbolically the sum of K nonidentical permutations of primed and unprimed quantities. Hence, S_N is formed from all the possible variants of the nonlocal observable $F_N=A_1A_2...A_N$ with different distributions of primes, and constitutes the determined dichotomic function

$$S_N(\lambda)=\pm 1 \quad (5.2.8)$$

with the average

$$\langle S_N\rangle_\rho=\int S_N(\lambda)\rho(\lambda)\mathrm{d}\lambda, \quad (5.2.9)$$

which, because, $\rho(\lambda)\geqslant 0$ and $\int\rho(\lambda)\mathrm{d}\lambda=1$ obeys the N-channel BI of the form

$$|\langle S_N\rangle_\rho|\leqslant 1. \quad (5.2.10)$$

We note that A_n and A_n' do not have to be dichotomic to satisfy (5.2.10). The sufficient condition is

$$|A_n|\leqslant 1, \quad |A_n'|\leqslant 1. \quad (5.2.11)$$

The Bell observable S_N can also be written more elegantly. We shall show this by Mermin's compact method[16] in a somewhat modified form. Consider the complex function

$$M_n\equiv A_n+iA_n', \quad \Pi_N\equiv\prod_{n=1}^{N}M_n. \quad (5.2.12)$$

It is readily verified that the iterational expressions for the Bell observable, given by (5.2.1) with the sign combinations given by (5.2.2) are identical with the following:

$$S_N=2^{-m}(\mathrm{Re}\,\Pi_N+\mathrm{Im}\,\Pi_N) \quad (5.2.13)$$

for the even $N=2m$ and

$$S_N=2^{-m}\,\mathrm{Re}\,\Pi_N \quad (5.2.14)$$

for the odd $N=2m+1$.

The function Π_N contains 2^N terms, which means that, according to (5.2.13), S_N includes $K=2^N$ terms; according to (5.2.14) the number of terms is $K=2^{N-1}$. For example, when $N=2$, 3, 4, 5 we have $K=4$, 4, 16, 16. On the other hand, in the iterational definition (5.2.1), $K_{N+1}=4K_N$. However, half of the terms in (5.2.1) cancel out for even N, and the remaining terms are equal in pairs and, as in (5.2.13) and (5.2.14), we have $K_{2m+1}=K_{2m}$.

Apart from compactness, the expressions given by (5.2.13) and (5.2.14) have the following further advantage: they can be readily used to calculate S_N, using the symbolic formula $\Pi_N=(A+iA')^N$ and Newton's binomial theorem, which was done in (5.2.4)–(5.2.7).

Next, we shall show how we can obtain BI directly from (5.2.13) and (5.2.14) by Mermin's argument.[16] According to (5.2.12), when (5.2.11) is satisfied, the extremal M_n and Π_N are given by

$$M_N=\pm 1\pm i=\pm\sqrt{2}\exp(iS_n\pi/4), \quad (5.2.15)$$

$$\Pi_N=\pm 2^{N/2}\exp(iS\pi/4), \quad (5.2.16)$$

where $S_n=\pm 1$, $S=\Sigma_{n=1}^{N}S_n=0$, $\pm 1...$. For even $N=2m$, we have $S=0$, ± 2, ± 4, ..., and

$$|\mathrm{Re}\,\Pi_N|=2^m,$$
$$|\mathrm{Im}\,\Pi_N|=0$$

or

$$|\mathrm{Re}\,\Pi_N|=0,$$
$$|\mathrm{Im}\,\Pi_N|=2^m. \quad (5.2.17)$$

For odd $N=2m+1$, we have $S=\pm 1$, ± 3, ..., and

$$|\mathrm{Re}\,\Pi_N|=|\mathrm{Im}\,\Pi_N|=(\sqrt{2})^{N-1}. \quad (5.2.18)$$

In all cases

$$S_N=\pm 1, \quad (5.2.19)$$

and we again arrive at the BI of (5.2.10).

5.3. Quantum theory of N-photon interference and violation of the Bell inequality

The aim of this Section is to show that the quantum average $\langle\psi|S_N|\psi\rangle\equiv\langle S_N\rangle_\psi$ can exceed unity for a certain combination of the phases $\{\alpha_n,\alpha_n'\}$.

To perform the optical variant of the experiment, we must prepare the N-photon state of the $2N$-mode field in the form

$$|\psi\rangle=\frac{1}{\sqrt{2}}\left(\prod_{n=1}^{N}a_{n1}^{+}+\prod_{n=1}^{N}a_{n2}^{+}\right)|0\rangle. \quad (5.3.1)$$

This is a natural generalization of the familiar states (3.4.1) and (4.5.1), which can be prepared, for example, with the help of $2N$-frequency parametric converters. The quantities $a_{n1,2}^{+}$ are operators representing the creation of photons in modes $n1$ and $n2$.

According (3.4.4) our observable is

$$A_n=\sigma_-^n e^{i\alpha_n}+\sigma_+^n e^{-i\alpha_n}=\sigma_x^n\cos\alpha_n+\sigma_y^n\sin\alpha_n, \quad (5.3.2)$$

where $\sigma_-^n=a_{n1}a_{n2}^{+}$, $\sigma_+^n=a_{n1}^{+}a_{n2}$.

We must now specify α_n and α_n'. Mermin[16] uses the restriction $\alpha_n=0$, $\alpha_n'=\pi/2$. Ardehali considers the variant $\alpha_n=\pi=4$, $\alpha_n'=3\pi/4$ (Ref. 18), and Mermin's results for even and odd N change places. We shall now examine a

more general case by imposing a less stringent condition on the phases [we recall that no phase restrictions were imposed in the derivation of the BI (5.2.10)]:

$$\alpha_n' = \alpha_n + \pi/2. \tag{5.3.3}$$

It then follows from (5.3.2) and (5.3.12) that

$$M_n = 2\sigma_+^n e^{-i\alpha_n}, \tag{5.3.4}$$

[(5.2.12)–(5.2.14) are now looked upon as operator relations in which $\mathrm{Re}\,\Pi_N \equiv (\Pi_N + \Pi_N^+)/2$, $\mathrm{Im}\,\Pi_N \equiv (\Pi_N - \Pi_N^+)/2i$ are, respectively, the Hermitian and non-Hermitian parts of Π_N, so that the average of the operator Π_N over the state (5.3.1) is simply [see (13)]

$$\langle \Pi_N \rangle_\psi = 2^N e^{-i\varphi} \langle \psi \mid \prod_{n=1}^{N} \left[\sigma_+^n (|10\rangle_n + |01\rangle_n) \right] / \sqrt{2}$$
$$= 2^{N-1} e^{-i\varphi}, \tag{5.3.5}$$

where

$$\varphi = \sum_{n=1}^{N} \alpha_n.$$

Thus, according to (5.2.13) and (5.2.14), we have

$$\langle S_N \rangle_\psi = 2^{m-1/2} \sin(\pi/4 - \varphi) \quad \text{for } N = 2m, \tag{5.3.6}$$

$$\langle S_N \rangle_\psi = 2^m \cos\varphi \quad \text{for } N = 2m+1. \tag{5.3.7}$$

We now take the resultant phase in (5.3.6) to be $\varphi = -\pi/4$, or 0 when the number of channels is odd [i.e., in (5.3.7)]. In both cases

$$\langle S_N \rangle_\psi = 2^{(N-1)/2} \tag{5.3.8}$$

and, if we use the BI (5.2.10), we obtain the final result given by (5.1). Its universal form—for both odd and even N—is different from Mermin's result[16] in which $\langle S_{2m} \rangle_\psi$ is smaller than our value by the factor $\sqrt{2}$. The reason for this is that the Bell observable is not defined in the same way: Mermin uses an expression such as (5.2.14) for any N; Roy and Singh[17] in their derivation of the more general relation (5.1) use instead of (5.3.1) a special state that is different for even and odd N.

Braunstein *et al.*[90] have shown that (5.3.8) gives the limiting value of the Bell observable in Mermin's form.[16] It is also the maximum value in our case.

The validity of (5.3.8) is also confirmed by the special cases of four- and five-channel systems. Since, in accordance with (5.3.1) and (5.3.2),

$$E_N \equiv \langle A_1 A_2 \ldots A_N \rangle_\psi = \cos\varphi, \quad \langle \Delta F_N^2 \rangle = \sin^2\varphi, \tag{5.3.9}$$

we find after taking the average of (5.2.6) and (5.2.7) that

$$\langle S_4 \rangle_\psi = \tfrac{1}{4}[\cos\varphi + 4\cos(\varphi + \pi/2) - 6\cos(\varphi + \pi)$$
$$- 4\cos(\varphi + 3\pi/2) + \cos(\varphi + 2\pi)], \tag{5.3.10}$$

$$\langle S_5 \rangle_\psi = \tfrac{1}{4}[\cos\varphi - 10\cos(\varphi + \pi) + 5\cos(\varphi + 2\pi)]. \tag{5.3.11}$$

The expression given by (5.3.10) resembles (3.1.11) for $N=2$. When $\varphi = -\pi/4$, each of the 16 terms in (5.2.6) and all the cosines in (5.3.10) (including their signs) are equal to $1/\sqrt{2}$, so that $\langle S_4 \rangle_\psi = 2^{3/2}$, in agreement with (5.3.8). As in the two-channel interferometer, complete correlations of the measurements do not, of course, occur.[18] However, it would be wrong to conclude that they are in principle impossible in the four-channel variant. Indeed, it was for $N=4$ that the GHZ paradox was first formulated[14] and the formulation relied on complete correlation (and anticorrelation). However, we have used only four correlators out of the possible sixteen, which for particular values of φ can be equal to ± 1; this obviously follows from (5.2.6) and (5.2.10) [see Sec. (5.5) for further details)]. The final conclusion that may be drawn from this discussion is that it is possible to generalize to arbitrary even N: although complete correlation (and anticorrelation) of the results of some measurements is possible, the conditions for this to happen are not identical with the requirement that $\varphi = -\pi/4$ which ensures that the Bell observable in the form given by (5.2.1) or (5.2.13) is a maximum.

The situation is different when the number of channels N is odd. For example, when $N=5$ and $\varphi=0$, all the 16 terms in (5.3.6) and the cosines in (5.3.11) (with allowance for the signs) are equal to unity. We then find that $\langle S_5 \rangle_\psi = 2^2$ is a maximum and we have complete correlation (and anticorrelation):[18] $E_N = \pm 1$ [if we take into account (5.3.9) and the primed A_n]. It is precisely under these conditions that we encounter the GHZ paradox for $N=3$ (see Sec. 4.3 for further details). Complete correlation is thus observed for all the phase combinations that we have employed (but with $\varphi=0$), and the observables $F_N = A_1 A_2 \ldots A_N$ (including primed A_n) do not fluctuate: in all samplings $F_N = \pm 1$. The last condition can be used in the experiment to establish reference values of the phases in each of the channels. In practice, the correlators E_N will differ from unity by the factor V (visibility) due to random coincidences (see Section 3.5). Consequently, the experimental value of the Bell observable is

$$\langle S_N \rangle_{\exp} = V_N 2^{(N-1)/2}, \tag{5.3.12}$$

and the necessary condition for the violation of the BI (5.2.10) is

$$V_N > 2^{(1-N)/2}. \tag{5.3.13}$$

Nevertheless, in two-photon experiments, the fulfillment of this type of condition is often 'faciliated' by introducing the corrected value $\langle \tilde{S}_2 \rangle = \langle S_2 \rangle_{\exp}/V_2$, so that the interference dependence 'automatically' leads to the violation of the BI.

5.4. Classical wave models

We shall now generalize some of the classical experiments considered in Secs. 3 and 4 to the case of arbitrary number N of observers. In particular, we shall verify that complete correlation or anticorrelation $E_N(\alpha_n) = \pm 1$ of the readings of N distant detectors whose sign is controlled by any of the N of the local parameters α_m is not itself specific to quantum models and is entirely amenable to classical treatment. The validity of the BI is then deter-

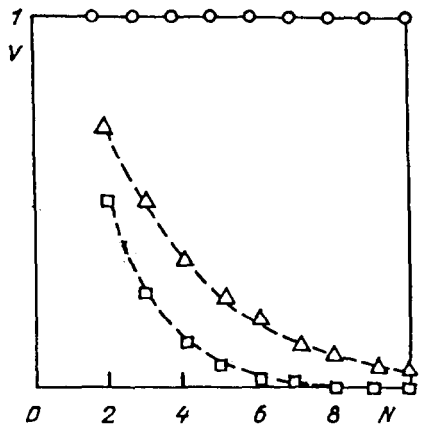

FIG. 7. Upper limits of visibility V above which the BI (triangles) and the predictions of the classical model (3.3.10) (squares) are found to fail. QT allows $V=1$ (circles).

mined not only by the visibility but also (or) by the character of the interference dependence of E_N on α_n.

We shall begin our discussion of classical models with a particular form of detection of continuous observables.

We return to the arrangements in Fig. 1 and consider an arbitrary number N of measuring channels instead of the two or three shown there. We also note that the parametric converters using N-frequency generators that emit $2N$ classical quasimonochromatic waves with constant intensities and slowly-drifting phases x_{nk}, $k=1$, 2 play the role of the hidden parameters λ_{nk}. These phases are related to the constant phases of the coherent pump waves φ_{01} and φ_{02} by the 'parametric' conditions (here and henceforth we shall put $x\equiv\lambda$)

$$\sum_{n=1}^{N}[\lambda_{n1}(t)-\lambda_{n2}(t)]=\varphi_{01}-\varphi_{02}. \tag{5.4.1}$$

Without loss of generality, we can put $\varphi_{01}=\varphi_{02}$ and

$$\sum_{n=1}^{N}\lambda_n=0,\quad \lambda_n\equiv\lambda_{n2}-\lambda_{n1}. \tag{5.4.2}$$

We note that the phase reference can also be taken to be a stable local homodyne if the N-channel scheme is organized by analogy with Fig. 5. A single parametric generator will then surfice. An additional phase α_n is then given to one of the signals in the measuring channels, and the two are mixed and rectified. Each observer measures the difference Δn_n between the rectified signals, which is given by the harmonic form (3.3.11). This achieves pure phase detection. Next, we use the measured quantities to form the correlator

$$E_N(\{\alpha_n\})=\langle A_1(\alpha_1,\lambda)A_2(\alpha_2,\lambda)...A_N(\alpha_N,\lambda)\rangle_\rho, \tag{5.4.3}$$

which in this case [for $\rho(\lambda)=\text{const}$], as in (3.3.12) and (4.4.2), describes the harmonic interference curve

$$E_N(\varphi)=V_N\cos\varphi. \tag{5.4.4}$$

We recall that the classical visibility V_N is subject to (3.3.10), i.e., a more stringent limitation then is necessary for BI [see (5.3.13)]. It follows that there exists an appreciable interval of values (Fig. 7)

$$2^{1-N}<V_N\leqslant 2^{(1-N)/2}, \tag{5.4.5}$$

that is consistent with HVT but is conflict with the specific classical wave model whose perculiar quantiative advantage, i.e., its smaller discrepancy from QT, is a consequence of this specificity.

The restriction on visibility, given by (3.3.10), is lifted in the discrete classical model presented here.

Let us now dichotomize the observables that appear in the phenomenological Bell theorem, using the sign function by analogy with (3.3.16) and (4.4.3):

$$A_n(\alpha_n,\lambda)\equiv\text{sign}\,\Delta n_n=\text{sign}\cos(\alpha_n+\lambda_n), \tag{5.4.6}$$

i.e., in each channel we have a bipolar telegraphic signal whose sign changes randomly and which has a correlation time τ_{coh}. As in the above case of continuous observables, all the signals are detected at time intervals $\Delta t\gg\tau_{\text{coh}}$. This variant yields N random dichotomic sequences $A_n(\alpha_n,t_i)=\pm 1$, $i=1$, 2, ..., $n=1$, 2,..., N.

The correlators of these sequences, for example (5.4.3), are determined by a distribution function which we assume to be factorizable:

$$\rho(\lambda)=\rho(\lambda_1,\lambda_2,...,\lambda_{N-1})=\prod_{n=1}^{N-1}\rho_n(\lambda_n). \tag{5.4.7}$$

This involves only $N-1$ random phases (of the independent hidden parameters) because one of the N phases is given by (5.4.2).

If we evaluate the average we find that

$$E_N(\alpha_1,\alpha_2,...,\alpha_N)$$
$$=\int\left[\prod_{n=1}^{N-1}A_n(\alpha_n,\lambda_n)\rho_n(\lambda_n)\right]A_N$$
$$\times\left(\alpha_N,-\sum_{n=1}^{N-1}\lambda_n\right)d\lambda_1...d\lambda_{N-1}. \tag{5.4.8}$$

When the phase distribution is uniform ($\rho_n=1/2\pi$) in the interval of 2π, complete correlation ($E_N=\pm 1$) can occur only in the two-channel interferometer (see Sec. 3.3). For $N=3$, we already have $|E_3|\leqslant 1/2$ [in both cases, the limitation on visibility in (3.3.10) is removed]. We cannot expect the return of complete correlation when $N\geqslant 3$. However, generally speaking, complete correlation is still possible in classical theory even for arbitrary N. Let us replace the uniform phase distribution with the inhomogeneous distribution

$$\rho_n(\lambda_n)=[\delta(\lambda_n)+\delta(\lambda_n-\pi+0)]/2,$$
$$-\pi<\lambda_n\leqslant\pi, \tag{5.4.9}$$

i.e., let us assume that the phase difference between the two original signals (or simply the phase of the signal in the case of the homodyne) in each of the channels changes by π at random instants of time. In other words, the ampli-

tudes of each input mode of the interferometer undergo independent and random changes of sign in pairs: $a_{nk} \to -a_{nk}$ (see Fig. 3c). We then find from (5.4.8) that (see the concluding part of Appendix IV)

$$E_N(\{\alpha_n\}) = \text{sign} \prod_{n=1}^{N} \cos(\alpha_n) = \pm 1. \tag{5.4.10}$$

The fact that the N-order moment is nonzero is a consequence of the unique relation between the phases in (5.4.2), and all the lower-order moments are all zero:

$$\langle A_n \rangle = \langle A_m A_n \rangle = \ldots = \langle A_1 A_2 \ldots A_{N-1} \rangle = 0, \quad m \neq n. \tag{5.4.11}$$

According to (5.4.10), $E_N = \pm 1$ as in the quantum model with odd N and $\varphi = 0$ or π [see (5.3.9)], but now E_N depends on each of the phases α_n individually and not merely on their sum

$$\varphi = \sum_{n=1}^{N} \alpha_n .$$

It is interesting to note that the sign of the correlator E_N can be determined by any of the N observers and that there are no intermediate correlation values (the only values are ± 1).

This last model refutes the commonly held view that the reason for BI violation is that correlations in quantum theory are stronger than those in classical theory. When the Cauchy-Schwartz inequality is violated [see (3.5.11)], this proposition is well founded, but in the case of BI, both the extremal values of the correlator and the specific form of the function $E_N(\alpha_n)$, i.e., the shape of the interference curve, are significant.

5.5. The GHZ paradox in the N-channel experiment

A detailed discussion of the GHZ paradox for three observers is given in Sec. 4.3. When $N=2$, the number of model parameters is insufficient and the paradox arises only at the cost of a considerable complication of the experimental method[22,23] (the essentially similar KS paradox may appear;[24,25] see Sec. 6.1). The case $N=4$ also deserves attention because it can be useful in elucidating the prospects for further increase in N. Moreover, it was precisely the four-channel thought experiment that was involved in the original GHZ proposal.[14]

Let us rewrite the correlator (5.3.9) in the form:

$$\begin{aligned} &\langle F_\varphi \rangle \equiv \langle A_\alpha B_\beta C_\gamma D_\delta \rangle = \cos\varphi, \\ &\langle \Delta F_\varphi^2 \rangle = \sin^2\varphi, \quad \varphi = \alpha + \beta + \gamma + \delta. \end{aligned} \tag{5.5.1}$$

We now perform four series of samplings in each of which complete correlation is realized ($\langle \Delta F^2 \rangle = 0$). We shall use the following sets of phases:

$$\begin{aligned} &(0,0,0,0,), \quad \left(\frac{\pi}{2}, \frac{\pi}{2}, 0, 0\right), \\ &\left(\frac{\pi}{2}, 0, \frac{\pi}{2}, 0\right), \quad \left(0, \frac{\pi}{2}, \frac{\pi}{2}, 0\right). \end{aligned} \tag{5.5.2}$$

We shall adopt the notation $A = A_0$, $A' = A_{\pi/2}$, $B = B_0$, $B' = B_{\pi/2}$, and so on, which ignores the sampling number (this 'simplification' is discussed in Sec. 4.3). According to (5.5.1), experiment gives the following results:

$$\begin{aligned} &F_0^{(1)} = ABCD = 1, \\ &F_\pi^{(2)} = A'B'CD = -1, \\ &F_\pi^{(3)} = A'BC'D = -1, \\ &F_\pi^{(4)} = AB'C'D = -1. \end{aligned} \tag{5.5.3}$$

We have omitted the averaging symbol because the observed F do not fluctuate, i.e., $\langle \Delta F^2 \rangle = 0$. By multiplying these expressions together, we obtain a contradiction of the form

$$Z = \prod_{m=1}^{4} F^{(m)} = (AA'BB'CC'D^2)^2 = -1. \tag{5.5.4}$$

We note that the phase δ of the fourth channel is not varied, i.e., we could do just as well without it.[15] This provides a trivial generalization of the formulation of the GHZ paradox to arbitrary N:

$$\begin{aligned} Z &= (A_1 A_1' A_2 A_2' A_3 A_3' A_4^2 \ldots A_N^2)^2 \\ &\equiv (A_1 A_1' A_2 A_2' A_3 A_3')^2 = -1. \end{aligned} \tag{5.5.5}$$

We thus see that nothing new is achieved by increasing the number of channels beyond 3. Moreover, the problem of ensuring that the set of hidden parameters λ falls into the same single-valuedness set Λ in all four series of experiments becomes even more acute, since the total number of basically possible outcomes of samplings increases as 2^{2N} and reliable demonstration of the inadequacy of HVT requires a larger number of realizations (see Sec. 4.3 for further details).

We also note that, within the framework of the above classical model with discrete phases, and despite the complete correlation, the GHZ paradox does not, of course, arise because the correlator (5.4.10) does not conform to condition (5.1.1).

6. OTHER EXAMPLES OF CONTRADICTIONS BETWEEN QUANTUM MECHANICS AND 'COMMON SENSE'

6.1. The Kochen–Specker theorem

We know that noncommuting operators cannot simultaneously assume determined numerical values. This follows formally from the fact that they do not have common eigenfunctions. Kochen and Specker (KS) have shown by considering a particular example that an algebraic contradiction of the form $+1 = -1$ arises when this rule is ignored.

Let us consider the simplest examples of proofs of the KS theorem, following Peres[24] and Mermin.[25] Suppose that we have two spin 1/2 particles. To be specific, we shall consider the experimental arrangement illustrated in Fig. 8 (see below, page 35). All the observable properties of the system are operator functions of the six 'basic' Pauli ma-

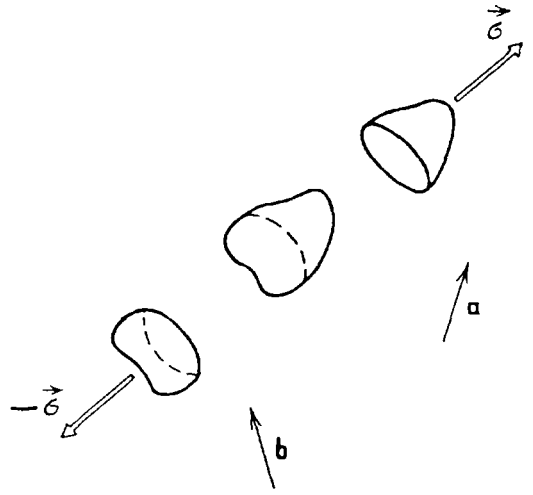

FIG. 8. The experiment proposed by Barut and Meystre.[93] $\vec{\sigma}$ and $-\vec{\sigma}$ are the angular momenta of the two halves of the original monolithic body. The projections of these vectors along **a** and **b** are measured.

trices $\sigma_\alpha^{(n)}$ ($n=1, 2$; $\alpha=x, y, z$) and the unity operator I. We shall put $\sigma_x^{(1)}=x_1$ and so on, and write down the basic relationships between these operators in the form

$$[x_n, y_n]=2x_n y_n=2iz_n,$$
$$[x_1, y_2]=[x_2, y_1]=0, \tag{6.1.1}$$
$$x_n^2=y_n^2=z_n^2=I.$$

We also define the operators

$$F_{\alpha\beta}=\sigma_\alpha^{(1)}\sigma_\beta^{(2)}, \tag{6.1.2}$$

which describe the correlation between the spin projections in the measurement channels.

We begin by assuming that the system is in a singlet state $|\psi\rangle$ of the form given by (II 6), in which case, according to (3.11),

$$\langle F_{xx}\rangle_\psi=\langle F_{yy}\rangle_\psi=\langle F_{zz}\rangle_\psi=-1, \tag{6.1.3}$$

where in the absence of random coincidences, the corresponding measurements will take place under the conditions of complete correlation with fluctuations having zero variance. We now form the product

$$F_{xy}F_{yx}\equiv x_1y_2y_1x_2=x_1y_1y_2x_2=z_1z_2=F_{zz}. \tag{6.1.4}$$

Since the operators F_{xy} and F_{yx} commute, a simultaneous measurement of the corresponding observables is in principle possible. Without specifying the details of this procedure, and allowing only for its potential realizability, we are entitled to expect from (6.1.4) and (6.1.3) that

$$\langle F_{xy}F_{yx}\rangle_\psi=-1 \tag{6.1.5}$$

even when the measurement results are completely correlated.

The eigenvalue spectrum of all the operators used in this Section consists of the two numbers ± 1, i.e., all the observables are dichotomic and experiments yield only these two values. We shall show later that (6.1.3) and (6.1.5) do not then fit into the HVT framework.

In the first series of samplings

$$\langle F_{xy}F_{yx}\rangle=x_1(\lambda_i^{(1)})x_2(\lambda_i^{(1)})y_1(\lambda_i^{(1)})y_2(\lambda_i^{(1)})=-1. \tag{6.1.6}$$

This is the HVT result and $\lambda_i^{(1)}$ is the set of hidden parameters of the particle source in the i-th sampling in the first series. The averaging symbols are omitted because the experimental result (6.1.5) is completely correlated with the single-valued outcome -1.

According to (6.1.3), the samplings in the second and third series should yield

$$\langle F_{xx}\rangle=x_1(\lambda_i^{(2)})x_2(\lambda_i^{(2)})=-1, \tag{6.1.7}$$

$$\langle F_{yy}\rangle=y_1(\lambda_i^{(3)})y_2(\lambda_i^{(3)})=-1. \tag{6.1.8}$$

As in the GHZ paradox, when the selections made in each series of samplings are sufficiently representative, we can find realizations for which the arguments $\lambda_i^{(n)}$ in (6.1.6)–(6.1.8) need not be taken into account, or they can be assumed to be the same throughout. The product of (6.1.7) and (6.1.8) is then identical in HVT with the left-hand side of (6.1.6), but is equal to $+1$, and we again encounter the $+1=-1$ contradiction.

All the above manipulations can readily be translated into the 'optical' language with the help of Appendices I and III. We note that, although this is not required in this case, measurement of the observables corresponding to $\sigma_z^{(n)}$ can be accomplished in accordance with Fig. 1a without the beamsplitters.

An interesting variant of the interference experiment, designed to demonstrate the $+1=-1$ contradiction, has been suggested by Hardy.[22] Its particular advantage as compared with the original GHZ proposal[14,15] lies in the use of the usual parametric source of two-photoin light, i.e., the experiment is feasible in the present state of development of modern technology. This also applies to the very recent paper by Bernstein *et al.*[23]

All the formulations of the $+1=-1$ paradox that we have examined so far, which included the GHZ paradox, relied on the preparation of singlet quantum states. The question is: is this condition necessary? The KS theorem gives a negative answer to this question: the original version of the theorem is valid for arbitrary states,[26] but involves the use of 117 spin one observed particles. Mermin[25] has succeeded in restricting number to 9. The following table is based on Peres' recipe:[24]

x_1	x_2	F_{xx}	I
y_2	y_1	F_{yy}	I
F_{xy}	F_{yx}	F_{zz}	$I.$
I	I	$-I$	

Around the table we show the products of operators in the corresponding columns and rows. Clearly, according to (6.1.1)

$$x_1x_2(x_1x_2)=x_1^2x_2^2=I, \tag{6.1.9}$$

$$y_2y_1(y_1y_2)=y_1^2y_2^2=I, \tag{6.1.10}$$

$$(x_1y_2)(y_1x_2)(z_1z_2)=x_1y_1z_1y_2x_2z_2=iz_1^2(-iz_2^2)=I, \tag{6.1.11}$$

$$x_1y_2(x_1y_2)=x_1^2y_2^2=I, \tag{6.1.12}$$

$$x_2y_1(y_1x_2)=y_1^2x_2^2=I, \tag{6.1.13}$$

$$(x_1x_2)(y_1y_2)(z_1z_2)=x_1y_1z_1x_2y_2z_2=(iz_1z_2)^2=-I. \tag{6.1.14}$$

The important point here is that the operators in the three-operator products in these equations the commute in pairs, so that they have common eigenstates in which they assume the value $+1$ or -1. It is therefore natural to replace the operators in the identities (6.1.9)–(6.1.14) with one of the two eigenvalues ± 1. These 'modified' operators will be indicated by the symbol [...]. The operator identities then assume the form of the following algebraic equations

$$\begin{gathered}[x_1][x_2][x_1x_2]=1,\ldots,\\ [x_1x_2][y_1y_2][z_1z_2]=-1,\end{gathered} \tag{6.1.15}$$

where the unity operators in (6.1.6) and (6.1.14) are replaced with 1. By multiplying together the left- and right-hand sides of these relations, we obtain the products of 18 eigenvalues:

$$[x_1][x_2][x_1x_2]\ldots[z_1z_2]=-1. \tag{6.1.16}$$

Each of the nine symbols [...] appears twice in the left-hand side of (6.1.16), so that

$$([x_1][x_2][x_1x_2]\ldots[z_1z_2])^2=-1, \tag{6.1.17}$$

i.e., $+1=-1$. The conclusion is that it isn't always possible to ascribe to the operators their eigenvalues. Claims such as 'algebraic proof of nonlocality of quantum theory' are sometimes encountered. We note in this connection that if in (6.1.16) we perform the reverse replacement of eigenvalues with operators (i.e., the remove the square brackets), the resulting operator identity involves noncommuting operators, for example, x_1 and y_1. Consequently, the contradictory equation (6.1.17) actually originates from the fact that we have ascribed eigenvalues to noncommunting operators.

The hidden parameters do not explicitly appear in the last version of the paradox, and locality is not emphasized although it is implied in the statement that observations of the first particle are independent of the orientation of the second detector (Stern–Gerlach analyzer) and the form of this independence determines whether we measure $\sigma_x^{(2)}$ or $\sigma_y^{(2)}$. We have already noted that this form of locality appears in both classical and quantum descriptions. It would seem that the logical escape from the contradiction (6.1.17) ought to be as follows: "for the above system, the QT formalism and the classical description with *a priori* values of the observables are mutually inconsistent." On the other hand, the interpretation of (6.1.17) as a manifestation of the nonlocality of QT seems unconvincing. We also note that the frequency used word 'contextual' also implies that the measuring devices in different channels depend on one another.

Formally, the GT and GHZ paradoxes follow from a property of the Lie algebra (6.1.1) of Pauli operators:

$$(\sigma_x\sigma_y)^2=(i\sigma_z)^2=-I. \tag{6.1.18}$$

However, any operator f that meets the condition $f^2=-I$ in a certain vector subspace of states is non-Hermitian, i.e., it cannot describe an observable variable and there is no meaning in ascribing to it a numerical value. Actually, the matrix elements of a Hermitian f are related by $f_{ij}=f_{ji}^*$ from which we have

$$(f^2)_{ii}=\sum_j |f_{ij}|^2\geqslant 0,$$

whereas it follows from $f^2=-I$ that $(f^2)_{ii}=-1$.

If we now turn to the practical aspect of the problem, we note that it is not obvious how we could experimentally demonstrate the contradiction defined by 6.1.17. One sampling cannot produce the values of all nine observables because there is no way of simultaneously measuring, say, $\sigma_x^{(1)}$ and $\sigma_y^{(1)}$. This means that there is no guarantee that all pairs of identical operators in (6.1.16) will have the same recorded value, since some of them may be obtained in different samplings. The paradox thus assumes a statistical character once again.

Having formulated the KS theorem for two spin 1/2 particles, we can readily generalize this to a large number of such particles. For example, for three particles we need ten observables and the corresponding operators.[25] In our notation, this follows immediately from (4.1.10). If we replace 1 with I in these formulas, and multiply the resulting operator identities from the left by $F^{(m)}$, we obtain

$$\begin{gathered}F^{(1)}A'BC=I,\quad F^{(2)}AB'C=I,\quad F^{(3)}ABC'=I,\\ F^{(4)}A'B'C'=I,\quad F^{(1)}F^{(2)}F^{(3)}F^{(4)}=-I.\end{gathered} \tag{6.1.19}$$

The first four relations actually have the form $(F^{(m)})^2=I$ and the last follows from (4.5.6). we also note that $[F^{(m)},F^{(m')}]=0$.

Multiplying together all five identities, we obtain an expression in which all the ten operators are encountered twice, and on the right we have $-I$. The subsequent steps repeat those that follow from (6.1.14).

It can be shown that the identities in (6.1.19) differ from the original identies in (4.1.10) by the fact that they are valid throughout the space of states of the system under consideration, which indicates that they are more general in character. This seems to be the only significant difference between the KS and GHZ paradoxes (a more detailed discussion is given in Ref. 25).

6.2. The Stapp contradiction

Stapp[73] has found a further interesting form of the contradiction between QT and the assumption of the *a priori* existence of observables independently of whether they are detected or not. This assumption is sometimes referred to as *contrafactual definiteness.*[73,74]

Stapp uses the frequency definition of averages, avoiding the explicit introduction of hidden parameters λ and simultaneous probabilities of noncommuting observables of the form $P(A,A')$. His variant of the corresponding experiment assumes both complete correlation between the readings obtained by two observers (see Fig. 1a) for $\varphi=0$ and $E=\cos\varphi=1$ and incomplete correlation ($|E|<1$). The paradox has common features with the Bell and GHZ theorems and occupies in effect and intermediate position between them. Stapp considers a experiment with two spins (Fig. 8), but we shall find it convenient in our presentation to use its optical analog (see Fig. 1a).

Consider four series of experiments ($m=1$, 2, 3, 4) with the following sets of phases:

$$\begin{aligned}&\alpha+\beta=0, \quad \alpha'+\beta=\pi/2,\\ &\alpha+\beta'=-\pi/4, \quad \alpha'+\beta'=\pi/4,\end{aligned} \tag{6.2.1}$$

For example,

$$\alpha=\beta=0, \quad \alpha'=\pi/2, \quad \beta'=-\pi/4.$$

For the state defined by (3.4.1), QT gives the following correlators:

$$\begin{aligned}&E^{(1)}\equiv\langle AB\rangle=1, \quad E^{(2)}\equiv\langle A'B\rangle=0,\\ &E^{(3)}\equiv\langle AB'\rangle=\frac{1}{\sqrt{2}}, \quad E^{(4)}\equiv\langle A'B'\rangle=\frac{1}{\sqrt{2}}.\end{aligned} \tag{6.2.2}$$

Since (6.2.2) is not in conflict with experimental data, we require that the classical average values should be the same.

As in the formulation of the KS and GHZ paradoxes, we ignore the dependence of the observables on the number of the sampling and of the experimental series. The significance of this was discussed in detail above. The assumption is theoretically justifiable, but in practice is satisfied by a limited selection from the set of realizations.

There is thus a series of quartets A_i', A_i', B_i, B_i', $i=1$, 2, ..., L that assume the values ± 1 and after averaging give the quantum results (6.6.2) as $L\to\infty$ e.g.,

$$\langle AB\rangle=\frac{1}{L}\sum_{i=1}^{L}A_iB_i.$$

Consequently, for complete correlation, $A_i=B_i$ and, replacing B_i with A_i in the second series of samplings, we have

$$\langle AA'\rangle=\frac{1}{L}\sum_{i=1}^{L}A_iB_i=0. \tag{6.2.3}$$

We recall once again that the symbol $\langle AA'\rangle$ does not have an operational meaning in QT.

Consider the following combination of observables:

$$T_i\equiv(A_i+A_i'-\sqrt{2}B_i')^2. \tag{6.2.4}$$

It can assume only three values, namely, $(2-\sqrt{2})^2\approx 0.34$, 2, and $(2+\sqrt{2})^2\approx 12$, i.e.,

$$\langle T\rangle\geqslant 0.34. \tag{6.2.5}$$

On the other hand, according to the definition given by (6.2.4), we have

$$T_i=4+2(A_iA_i'-\sqrt{2}A_iB_i'-\sqrt{2}A_i'B_i'), \tag{6.2.6}$$

where we have used the fact that $A_i^2=A_i'^2=B_i'^2=1$.

However, if we average (6.2.6) and use (6.2.2) and (6.2.3), we obtain

$$\langle T\rangle=4-2\sqrt{2}(\langle AB'\rangle+\langle A'B'\rangle)=0, \tag{6.2.7}$$

which is not consistent with (6.2.5).

It follows that the classical pair moments cannot assume values from the quantum set defined by (6.2.2).

6.3. Contradiction based on the Cauchy–Schwartz inequality

The paradox formulated below is remarkable in that it gives rise to enormous discrepancies between the classical and quantum theories (by eight orders of magnitude!), but it does not have the universality typical of the Bell, KS, and GHZ theorems because it is associated with a particular model of an experiment. The model is simple. Suppose that each of the parametrically created pair photons is directly recorded by its own detector, i.e., in the scheme of Fig. 1a there is only one piezocrystal and two photodetectors connected in coincidence.

The two-photon state of the two-mode field is prepared by a parametric converter and is described by the vector given by (III6). We now rewrite the formula in the somewhat more general form

$$|\psi\rangle=C_{00}|0\rangle_a|0\rangle_b+C_{11}|1\rangle_a|1\rangle_b, \tag{6.3.1}$$

so that the lower-order moments are

$$\langle ab\rangle_\psi=C_{00}^*C_{11}, \tag{6.3.2}$$

$$\langle n_a\rangle_\psi=\langle n_b\rangle_\psi=\langle n_an_b\rangle_\psi=|C_{11}|^2. \tag{6.3.3}$$

The quantity $\langle n_an_b\rangle$ gives the general probability of coincidence between photon counts, of which a fraction proportional to $\langle n_a\rangle\langle n_b\rangle$ does not depend on the time de-

lay and consist of random coincidences that provide the background. For example, the maximum of the correlation function

$$G(\Delta t) \equiv \langle n_a(t_a) n_b(t_b) \rangle, \quad \Delta t = t_a - t_b \tag{6.3.4}$$

corresponds to the moment $\langle n_a n_b \rangle$ at $\Delta t = 0$ and its wings (pedestal) correspond to the products $\langle n_a \rangle \langle n_b \rangle = G(\infty)$, where t_a and t_b at the times taken by photons between the source and the detectors in the signal and idle channels, respectively.

For single-mode detectors, the photon count rate in each channel is given by (3.5.6):

$$R_\nu = \eta'_\nu \langle n_\nu \rangle / T, \tag{6.3.5}$$

where $\nu = a, b$.

Similarly, the total coincidence rate is

$$R_c \eta'_a \eta'_b \langle n_a n_b \rangle / T, \tag{6.3.6}$$

and the random coincidence rate is

$$R_{\rm acc} = \eta'_a \eta'_b \langle n_a \rangle \langle n_b \rangle / T = R_a R_b T. \tag{6.3.7}$$

The photon bunching parameter is thus given by

$$g \equiv \frac{G(0)}{G(\Delta t \to \infty)} = \frac{\langle n_a n_b \rangle}{\langle n_a \rangle \langle n_b \rangle} = \frac{R_c}{R_{\rm acc}} = 1 + \mu^2. \tag{6.3.8}$$

where the ratio of true to random coincidences is

$$\mu^2 \equiv \frac{R_c - R_{\rm acc}}{R_{\rm acc}} = \left| \frac{C_{00}}{C_{11}} \right|^2 = \frac{|\langle ab \rangle|^2}{\langle n_a n_b \rangle}. \tag{6.3.9}$$

This makes use of the normalization condition

$$\langle \psi | \psi \rangle \equiv |C_{00}|^2 + |C_{11}|^2 = 1. \tag{6.3.10}$$

In classical theory, the coefficient μ is subject to the restriction[154]

$$|\mu_{\rm class}| \leqslant 1, \tag{6.3.11}$$

which follows from from the Cauchy–Schwartz inequality.[50]

The fact that μ does not have to obey (6.3.11) was discovered in Ref. 163 and was verified experimentally by Burnham and Weinberg.[162] For parametric scattering, it is found that, typically, $|\mu|^2 \sim 10^8$. This kind of difference between the maximum of the correlation function and the background is evidence that light cannot not be described by a classical wave in this experiment.

7. CONCLUSION

The flood of publications concerned with the questions discussed above shows no signs of abating, and it is likely that this review, in which we have tried to present the most interesting ideas and experiments published up to and including 1992, will probably be out of date by the time it is published. Indeed, some interesting publications have only just appeared.[186–203]

Nevertheless, we hope that our review will help readers to acquire an appreciation of the current literature, and that the detailed analysis and clear classical models presented above will provide a basis for new ideas.

We have tried, on the one hand, to attract the attention of the 'silent majority' of physicists to quantum paradoxes that can be demonstrated by optical methods, whilst, on the other hand, we have attempted to to sharpen up the meaning of the widely used word 'nonlocality'. The optical interference effects discused above are, in our view, no more nonlocal than the effects described by classical statistical optics and the semiclassical theory of photodetection. In the Heisenberg approach, the propagation of light in interferometers and other linear optical devices is described in terms of the essentially classical Green's functions (propagators), so that the quantum specificity reduces merely to unusual statistical properties (violation of the Cauchy–Schwartz inequality by mode amplitudes) of the 'nonclassical' light source used at entry to the interferometer.

The above models with phase-coded classical noise waves generated by parametric devices (Secs. 3.3, 4.4, and 5.4) show that complete correlation of the readings of N distant detectors, $E(\alpha_1, \alpha_2, \alpha_3, ..., \alpha_N) = \pm 1$, whose sign is determined by any of the parameters α_i, is not the exclusive prerogative of quantum models, contrary to the widely held view.

Although studies of the EPR-Bell paradox are essentially methodological and interpretational ('metaphysical'[10]) in character, history of physics shows that they can lead to important practical consequences, e.g., the development of new interferometers and data-transmission techniques.[132–135] Studies in quantum photometry[136–138] and quantum cryptography[128–131], using two-photon light, can also be of practical interest.

Finally, the continuing hunt for logical loopholes that has been casting serious doubts on experimental demonstrations of the incompatibility of 'local realism' and quantum formalism (cf., for example, the crytical analysis by Santos[44,45]) has stimulated the emergence of new ideas and new avenues for experimental research. It has also been beneficial in reducing excessive conformity and complacency that has impeded creativity.

It is my pleasant duty to thank Alevtina Prokhorovna Krylova for her unfailing and invaluable help in this research.

APPENDICES

I. Spins, photons, and phases

We shall now consider a set of two-mode single-photon states in which a photon belongs to one of the two modes with equal frequencies:

$$|\psi\rangle = C_{01}|01\rangle + C_{10}|10\rangle, \quad |C_{01}|^2 + |C_{10}|^2 = 1. \tag{I1}$$

These states form a Hilbert space that is isomorphous with the space of states of a spin 1/2 particle or, generally, a two-level quantum system. We know that this space maps on to a sphere of unit radius (Bloch or Poincaré sphere). To link the photons to the spins, we introduce the following notation:

$$\sigma_- \equiv a_1 a_2^+ = (\sigma_x - i\sigma_y)/2,$$
$$\sigma_+ \equiv a_1^+ a_2 = (\sigma_x i\sigma_y)2,$$
$$\sigma_x \equiv \sigma_- + \sigma_+, \quad \sigma_y \equiv i(\sigma_- - \sigma_+), \tag{I2}$$
$$\sigma_z \equiv n_1 - n_2 = [\sigma_+, \sigma_-],$$

so that

$$\sigma^2 \equiv \sigma_x^2 + \sigma_y^2 + \sigma_z^2 = 4j(j+1)$$

where $j \equiv (n_1 + n_2)/2$. The operators $\sigma_{x,y,z}$ correspond to Pauli operators in the space (I1).

It follows from these relations that the operator $\sigma_- = (\sigma_+)^+$ transfers a photon from the first mode to the second and σ_+ performs the reverse operation:

$$\sigma_- |10\rangle = |01\rangle, \quad \sigma_- |01\rangle = 0,$$
$$\sigma_+ |01\rangle = |10\rangle, \quad \sigma_+ |10\rangle = 0. \tag{I3}$$

From (I 1) and (I 2),

$$\sigma_- |\psi\rangle = C_{10} |01\rangle,$$
$$\sigma_+ |\psi\rangle = C_{01} |10\rangle,$$
$$\sigma_x |\psi\rangle = C_{01} |10\rangle + C_{10} |01\rangle,$$
$$\sigma_y |\psi\rangle = -iC_{01} |10\rangle + iC_{10} |01\rangle, \tag{I4}$$
$$\sigma_z |\psi\rangle = C_{10} |10\rangle - C_{01} |01\rangle,$$
$$2j |\psi\rangle = |\psi\rangle.$$

Repeated application of the operators yields

$$\sigma_+^2 |\psi\rangle = \sigma_-^2 |\psi\rangle = 0, \quad \sigma_- \sigma_+ |\psi\rangle = C_{01} |01\rangle,$$
$$\sigma_+ \sigma_- |\psi\rangle = C_{10} |10\rangle, \tag{I5}$$
$$\sigma_x^2 |\psi\rangle = \sigma_y^2 |\psi\rangle = \sigma_z^2 |\psi\rangle = |\psi\rangle.$$

Thus (I1) is an eigenstate of the operators $2j, \sigma_x^2, \sigma_y^2, \sigma_z^2$ with eigenvalue $+1$ and of the operators σ_+^2 and σ_-^2 with eigenvalue 0. Consequently, in this single-photon Hilbert space, these operators are scalars

$$2j = \sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I,$$
$$\sigma_+^2 = \sigma_-^2 = 0. \tag{I6}$$

From (I2) we also have the following relations typical of Pauli matrices and, in general, the Lie algebra of SU(2):

$$[\sigma_1, \sigma_2] = 2i\sigma_3, \tag{I7}$$

where the subscripts 1, 2, 3 are obtained from x, y, z by even permutation.

The transformation matrices for the state vectors of the field (I1) that correspond to linear optical elements (beamsplitters, interferometers, polarizers, etc.) can be given in the form of linear combinations of Pauli matrices, and this also applies to quantities observed with two photon counters. This is an example of a general rule: observables play a two-fold role—they can be measured and they are generators of transformations. For example, according to (3.4.4), the observed difference between the readings of two photon counters at the exit of the beamsplitter is described by the operator

$$A_\alpha \equiv n_+ - n_-$$
$$= \sigma_- e^{i\alpha} + \sigma_+ e^{-i\alpha}$$
$$= \sigma_x \cos\alpha + \sigma_y \sin\alpha$$
$$= \vec{\sigma}\mathbf{n}_\alpha, \tag{I8}$$

where $\mathbf{n}_\alpha = (\cos\alpha, \sin\alpha)$ is a unit vector. For a spin 1/2 particle, this observable corresponds to the spin projection along $\mathbf{n}_\alpha$ in the xOy plane, i.e., $\mathbf{n}_\alpha$ gives the orientation of the magnet in the Stern-Gerlach experiment. If, on the other hand, we are concerned with the single-photon field, the vector $\mathbf{n}_\alpha$ defines the phase shift α between the modes incident on the beamsplitter.

In the general case, these modes are distinguished both by the direction of the wave vector and the polarization. When the modes belong to a single plane wave, the vector $\langle\vec{\sigma}\rangle$ is related to the Stokes parameters, and an arbitrary linear transformation of the state (I1) is related to the Jones matrices.[33] The beamsplitter and the phase delay are then replaced by a rotator that rotates the plane of polarization through the angle 2α. If, on the other hand, the states $|10\rangle$, $|01\rangle$ are referred to two modes with opposite circular polarizations, the average $\langle\sigma_z\rangle = |C_{10}|^2 - |C_{01}|^2$ can be interpreted simply as the angular momentum transported by the plane wave.

The definition given by (I2) enables us to use the vector $\vec{\sigma}$ (apart from an arbitrary phase) to map any state of the two-mode field with given energy (classical or quantum) on to a sphere. In classical theory this takes the form

$$\sigma_x = 2\,\mathrm{Re}(a_1^* a_2), \quad \sigma_y = 2\,\mathrm{Im}(a_1^* a_2),$$
$$\sigma_z = |a_1|^2 - |a_2|^2, \tag{I9}$$
$$(\sigma_x^2 + \sigma_y^2 + \sigma_z^2)^{1/2} = |a_1|^2 + |a_2|^2 = \mathrm{const}.$$

Let us now return to quantum theory and continue our study of the properties of A_α. We can readily show, using (I3) and (I8), that the vectors

$$|\psi_\alpha^+\rangle = (e^{-i\alpha/2}|10\rangle + e^{i\alpha/2}|01\rangle)/\sqrt{2},$$
$$|\psi_\alpha^-\rangle = (e^{-i\alpha/2}|10\rangle - e^{i\alpha/2}|01\rangle)/\sqrt{2}, \tag{I10}$$

are special cases of (I1), i.e., the eigenstates of A_α:

$$A_\alpha |\psi_\alpha^+\rangle = |\psi_\alpha^+\rangle, \quad A_\alpha |\psi_\alpha^-\rangle = -|\psi_\alpha^-\rangle. \tag{I11}$$

Assuming that $\alpha = 0$ and $\alpha' = \pi/2$, we find the eigenvectors for σ_x, σ_y:

$$|\psi_x^+\rangle = (|10\rangle + |01\rangle)/\sqrt{2},$$
$$|\psi_y^+\rangle = e^{-i\pi/4}(|10\rangle + i|01\rangle)/\sqrt{2},$$
$$|\psi_x^-\rangle = (|10\rangle - |01\rangle)/\sqrt{2},$$
$$|\psi_y^-\rangle = e^{-i\pi/4}(|10\rangle - i|01\rangle)/\sqrt{2}. \tag{I12}$$

Any two-mode state such as (I1) can always be written as a linear combination of two other states of the same type, e.g., $|\psi_x^-\rangle$ and $|\psi_y^-\rangle$.

We emphasize that $A_\alpha, A_{\alpha'}$ do not commute for $\alpha \neq \alpha'$, so that they do not have common eigenvectors.

We now introduce the Hermitian coordinate and momentum operators (they are identical, apart from the factor $\sqrt{2}$, with the quadrature components X, Y mentioned in Sec. 3.6):

$$q_k=(a_k+a_k^+)/\sqrt{2}, \quad p_k=(a_k-a_k^+)/i\sqrt{2}, \tag{I13}$$

and the observable A_α assumes the form

$$A_\alpha=(q_1q_2+p_1p_2)\cos\alpha+(q_1p_2-p_1q_2)\sin\alpha. \tag{I14}$$

There is one other possible interpretation of A_α, based on the phase operators $E^\pm$ and the phase difference between two oscillators C_{12}, S_{12} (Ref. 159)

$$E_k^+=a_k^+(n_k+1)^{-1/2},$$

$$E_k^-=(n_k+1)^{-1/2}a_k,$$

$$C_{12}=(E_1^-E_2^++E_1^+E_2^-)/2,$$

$$S_{12}=(E_1^-E_2^+-E_1^+E_2^-/i2. \tag{I15}$$

Their classical analogs are

$$E^-\to a/|a|=e^{-i\phi}, \quad E^+\to e^{i\phi},$$

$$C_{12}\to\cos(\phi_1-\phi_2)S_{12}\to\sin(\phi_1-\phi_2). \tag{I16}$$

In the case of the two-mode Fock state $|n_1n_2\rangle$, the operators $E_1^\pm E_2^\mp$ differ from $\sigma_\pm$ only by a numerical factor:

$$E_1^-E_2^+=\sigma_+[n_1(n_2+1)]^{-1/2},$$

$$E_1^+E_2^-=\sigma_-[(n_1+1)n_2]. \tag{I17}$$

According to (I1) and (I4),

$$C_{12}|\psi\rangle=(C_{01}|10\rangle+C_{10}|01\rangle)/2=\tfrac{1}{2}\sigma_x|\psi\rangle,$$

$$S_{12}|\psi\rangle=i(C_{01}|10\rangle-C_{10}|01\rangle)/2=-\tfrac{1}{2}\sigma_y|\psi\rangle. \tag{I18}$$

In the single-photon space (I1), our observables are therefore identical with the phase difference operators multiplied by 2:

$$A_0=a_1a_2^++a_1^+a_2=\sigma_x=2C_{12},$$

$$A_{\pi/2}=i(a_1a_2^+-a_1^+a_2)=\sigma_y=-2S_{12}. \tag{I19}$$

Thus, in a single-photon two-mode state, the operators C_{12}, S_{12} have eigenvalues $\pm 1/2$ (Ref. 159), which agrees with the eigenvalues ± 1 of the operator A_α. We recall that the phase of single-mode Fock states is undefined. The relations given by (I19) are usefully compared with the definition of observables given by (3.3.11) in the classical stochastic model. This shws that anticorrelation between the counts produced by two detectors at exit from the beamsplitter in one of the interferometer channels in the case of a two-mode single-photon state (i.e., a photon recorded by one detector, say, D_+^a cannot be accompanied by a simultaneous count in another, D_-^a) can be looked upon as a demonstration of the quantization of the phase difference. In other words, each channel of the interferometer is a phase-difference meter and, specifically, measures C_{12}, S_{12} (in radio-frequency practice, this is simply a phase detector). We note that the definition of the phase operator and its measurement in quantum optics have recently attracted considerable attention.[180–183]

The above discussion can be extended to an arbitrary N-photon two-mode state that corresponds to the state of a particle with spin $j=N/2$ (Refs. 33 and 39). It is interesting to note that the correspondence between the transformation properties of the two-mode field and the group SU(2) is not specific to the quantum field because the classical N-order moments $a_1^na_2^{N-n}/n^{1/2}(N-n)^{1/2}$ where $n=0,1,2,\ldots, N$ also form a basis of this group with dimension $N+1$ (Refs. 33, 184).

II. Model employing two particles with anticorrelated angular momenta

Let us now consider a thought experiment that essentially repeats Bell's model[3,15] but incorporates the special feature introduced by Barut and Meystre.[93] Suppose that some material body with zero initial angular momentum splits into two parts whose angular momenta are equal and opposite in accordance with the conservation rules. We shall denote them by $\vec{\sigma}$ and $-\vec{\sigma}$, respectively (Fig. 8). We shall measure the projections of the angular momenta of the two particles: one along **a** and the other along **b** where $|\mathbf{a}|=|\mathbf{b}|=1$. Measurement then yields

$$A\equiv\vec{\sigma}a=\sigma\sin\theta\cdot\cos\phi,$$

$$B\equiv-\vec{\sigma}b=-\sigma\sin\theta\cdot\cos(\phi-\varphi), \tag{II1}$$

where $\sigma=|\vec{\sigma}|$, φ is the angle between the vectors **a**, **b**, and ϕ and θ are the spherical (polar) coordinates, such that the longitude is measured from the direction of **a** and the polar angle θ from the normal to the plane containing **a** and **b**.

Multiple repetition of the experiment for a constant magnitude and uniform angular distribution of the angular momentum $\vec{\sigma}$ gives

$$\langle A\rangle=\langle B\rangle=0, \tag{II2a}$$

$$\langle A^2\rangle=\frac{\sigma^2}{4\pi}\int_0^{2\pi}\cos^2\phi\,\mathrm{d}\phi\int_0^\pi\sin^3\theta\,\mathrm{d}\theta=\frac{\sigma^2}{3}=\langle B^2\rangle. \tag{II3a}$$

This integral actually represents the locus of the end-point of the vector $\vec{\sigma}$ on the surface of a sphere of radius σ and the continuous evaluation of the projection of this vector along **a** (or **b**). It is precisely for this reason that the integral in (II3a) acquires an 'extra' sine. Indeed, if this projection were identically equal to unity, then the integral

$$\int_0^{2\pi}\mathrm{d}\phi\int_0^\pi\sin\theta\,\mathrm{d}\theta=4\pi$$

would be the area of a sphere of unit radius.

Let us define the correlation between the quantities that are being measured, normalized to σ^2:

$$E_\varphi\equiv\frac{\langle AB\rangle}{\sigma^2}=-V\cos\varphi, \quad V=\tfrac{1}{3}. \tag{II4a}$$

The normalization that we have used corresponds to the condition for Bell's theorem (3.2.9) because $|A|/\sigma \leqslant 1$ and $|B/\sigma \leqslant 1$.

The physical significance of the number 3 in the denominator of (II3a) and of the visibility V in (II4a) is clear: the mean squares of the three components of angular momentum are equal, i.e., $\langle \sigma_x^2 \rangle = \langle \sigma_y^2 \rangle = \langle \sigma_z^2 \rangle = \sigma^2/3$, because the model is isotropic. It is precisely because it is isotropic that the results given by (II2a)–(II4a) do not depend on the absolute orientation of the measuring devices in space, but only on the angle between them [in (II4a)].

The classical limit for the visibility therefore takes the following form in this case:

$$V \leqslant 1/3, \tag{II5a}$$

because the nonideality of detectors can only reduce this limit. We recall that, in the classical stochastic model of the interferometer, considered at the beginning of Sec. 3.3, the limit was $V \leqslant 1/2$ [see (3.3.8)].

Bell's original model[3,15] is very close to the experiment just described. The only difference is in the algorithm used to process the experimental data: the sign function was used by Bell to transform A and B into dichotomic variables by analogy with (3.3.16). The dependence on φ in the correlation function then takes the sawtooth form (Fig. 3b).

We now turn to the quantum analysis of this experiment. We assume that the particles travelling in opposite directions have anticorrelated spins $\langle j \rangle = \pm 1/2$. Their quantized state will be defined by the vector

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|+\rangle_a |-\rangle_b - |-\rangle_a |+\rangle_b), \tag{II6}$$

where

$$\sigma_z^a |+\rangle_a = |+\rangle_a, \quad \sigma_z^a |-\rangle_a = -|-\rangle_a,$$
$$\sigma_z^b |+\rangle_b = |+\rangle_b, \quad \sigma_z^b |-\rangle_b = -|-\rangle_b, \tag{II7}$$

i.e., the vectors $|\pm\rangle_{a,b}$ are the eigenvectors of the Pauli operators $\sigma_z^{a,b}$ with eigenvalues ± 1 (here and henceforth, σ_x, σ_y, σ_z represent the Pauli matrices).

The disposition of the measuring devices in space will be described in its general form by

$$\mathbf{a} \equiv \{a_x, a_y, a_z\} \equiv \{\sin\theta_a \cdot \cos\phi_\alpha, \sin\theta_a \cdot \sin\phi_\alpha, \cos\theta_a\} \tag{II8}$$

and similarly for $\mathbf{b}$.

The measured quantities are then

$$A = a_x \sigma_x^a + a_y \sigma_y^a + a_z \sigma_z^a,$$
$$B = b_x \sigma_x^b + b_y \sigma_y^b + b_z \sigma_z^b. \tag{II9}$$

Repetition of these measurements yields dichotomic sequences consisting of ± 1. The conditions of Bell's theorem are thus met.

When correlators such as (II2a)–(II4a) are evaluated, we need relations of the form[158]

$$\sigma_x |+\rangle = |-\rangle, \quad \sigma_x |-\rangle = |+\rangle,$$
$$\sigma_y |+\rangle = i|-\rangle, \quad \sigma_y |-\rangle = -i|+\rangle; \tag{II10}$$

where subscripts a and b are temporarily discarded.

By averaging over the states (II6) for all the possible combinations of Pauli operators, we therefore obtain

$$\langle \sigma_\alpha^{a,b} \rangle_\psi = \langle \sigma_\alpha^{a,b} \sigma_\beta^{a,b} \rangle = 0 \quad \text{for } \alpha \neq \beta,$$
$$\langle (\sigma_\alpha^a)^2 \rangle_\psi = \langle (\sigma_\alpha^b)^2 \rangle_\psi = 1,$$
$$\langle \sigma_\alpha^a \sigma_\alpha^b \rangle_\psi = -1, \tag{II11}$$

where α, $\beta = x, y, z$.

The required correlators are now readily determined:

$$\langle A \rangle_\psi = \langle B \rangle_\psi = 0, \tag{II2b}$$

$$\langle A^2 \rangle_\psi = \langle B^2 \rangle_\psi = |\mathbf{a}|^2 = |\mathbf{b}|^2 = 1, \tag{II3b}$$

$$E_\varphi \equiv \langle AB \rangle_\psi = -\mathbf{ab} = -\cos\varphi, \tag{II4b}$$

where, as before, φ is the angle between the measuring devices.

It is clear that the final results again depend only on the mutual orientation of the measuring systems. The quantum model (like the classical model) is isotropic. However, in conflict with (II5a), the visibility in QT is $V = 1$. Consequently, BI such as (3.2.8) may be violated.

It is interesting to note that the correlation factor [cf. (3.3.15)]

$$\Gamma_\varphi = \frac{\langle AB \rangle}{(\langle A^2 \rangle \langle B^2 \rangle)^{1/2}} = -\cos\varphi, \tag{II12}$$

is the same in both quantum and classical models. The clasical model does not then meet condition (3.2.9) of the Bell theorem because the measured relative quantities $A/\langle A^2 \rangle^{1/2}$, $B/\langle B^2 \rangle^{1/2}$ may be numerically greater than 1.

The following observations are appropriate in view of the closeness of the above thought experiment, on the one hand, and the Bell theorem and the EPR paradox, on the other.[1-3] The point is that particular misunderstandings may arise in the analysis of the proccess of measurement of the parameters of correlated moving particles. Actually, whilst for a single particle we cannot measure two observables described by noncommuting operators (e.g., σ_x and σ_y), we find that, for two particles, correlation appears to offer us the possibility of sidestepping this difficulty: for one of them we might measure σ_x^a in, say, channel A and for the other we might measure σ_y^b, and since $\sigma_x^a = -\sigma_x^b$, $\sigma_y^a = -\sigma_y^b$, we might obtain the entire necessary information (for the sake of brevity, we identify observables with their operators). In actual fact, this is not a sensible scenario if only because σ_x^a and σ_y^b are uncorrelated: $\langle \sigma_x^a \sigma_y^b \rangle = \langle \sigma_x^a \rangle \langle \sigma_y^b \rangle = 0$ [see (II11)], i.e., such measurements on correlated particles are equivalent to measurements on totally uncorrelated particles, and this property is not exclusive to QT: in the classical model, σ_x^a and σ_y^b have to measured with detectors at an angle $\varphi = \pi/2$ to one another. According to (II4a), however, we then have $\langle AB \rangle = \langle A \rangle \langle B \rangle = 0$. Such paradoxes do not therefore arise in this model.

III. Radiation from a parametric frequency converter in an interferometer

The simplest model of two-mode parametric amplification in the field of a steady classical plane monochromatic pump is based on the Hamiltonian[33]

$$H=i\hbar\Gamma(a^+b^+-ab), \tag{III1}$$

where $a^+(a)$ and $b^+(b)$ are the photon creation (annihilation) operators in the signal and idle modes and Γ is the gain parameter proportional to the nonlinearity $\chi^{(2)}$, its length z, and the pump amplitude. For simplicity, we suppose that Γ is real, which is equivalent to specifying a particular pump phase angle.

From (III1) we obtain the Heisenberg equations of motion

$$\frac{\mathrm{d}a}{\mathrm{d}t}=\Gamma b^+, \quad \frac{\mathrm{d}b}{\mathrm{d}t}=\Gamma a^+, \tag{III2}$$

that describe the 'mixing' of the creation and annihilation operators. They can also be used to describe the classical parametric amplifier if we replace the operators a^+, b^+ with negative-frequency slowly-varying mode amplitudes a^*, b^* in dimensionless units, and a, b with positive-frequency amplitudes. In an actual stationary optical experiment, the process evolves in space but not in time, and t is replaced with z/c (see Ref. 179 for a justification of this).

In the Schrödinger representation, it is the field state vector (and not the operators) that evolves:

$$i\hbar\mathrm{d}|\psi\rangle/\mathrm{d}t=H|\psi\rangle. \tag{III3}$$

The solution of this equation can be written in the form

$$|\psi(\tau)\rangle=U(\tau)|\psi_0\rangle, \quad U(\tau)=e^{\tau(a^+b^+-ab)}, \tag{III4}$$

where $|\psi_0\rangle$ is the initial state of the field and $\tau=\Gamma t$. The evolution operator U is now called the *squeezing operator* because, in parallel with the generation of biphotons, we have the suppression of the quantum fluctuations in one of the quadrature components of the resultant field of the signal and idle modes.[168–170]

The equivalent solution of (III2) in the Heisenberg representation takes the form of the Bodolyubov unitary transformation

$$a=U^+a_0U=ua_0+vb_0^+, \quad b=ub_0+va_0^+ \tag{III5}$$

where $u=\cosh\tau$ and $v=\sinh\tau$.

Suppose that the initial state of the field is the vacuum field $|\psi_0\rangle=|0\rangle$, so that in first-order perturbation theory in τ we find from (III4) that

$$|\psi\rangle=(I+\tau a^+b^+)|0\rangle=|0\rangle+\tau|1\rangle_a|1\rangle_b, \tag{III6}$$

where $|1\rangle_a|1\rangle_b\equiv|1\rangle_a\otimes|1\rangle_b$ is the state of field with the simultaneous presence of one signal and one idle photon.

Since τ is small ($\ll 1$), the specific contribution of vacuum to (III6) is very considerable. It is precisely the vacuum component of $|\psi\rangle$ that is responsible for the nonzero nonstationary 'anomalous' correlator that emerges when when the 'fast' time dependence is taken into account and takes the form

$$\langle a(t)b(t)\rangle=\tau\langle 0|ab|1\rangle_a|1\rangle_b e^{-i(\omega_a+\omega_b)t}=\tau e^{-i\omega_0 t}, \tag{III7}$$

where $\omega_0=\omega_a+\omega_b$ is the pump frequency. The presence of this nonstationary moment is seen directly in homodyne detection [cf., for example, (3.7.1)]. If, on the other hand, we use conventional detectors of photon coincidences that are not sensitive to the vacuum component, we may assume that

$$|\psi\rangle=a^+b^+|0\rangle\equiv|1\rangle_a|1\rangle_b. \tag{III8}$$

The next step in implementing interference experiments on BI verification is to distribute the signal and idle photons over two further modes because we need four-mode intensity interference.[33,34] The simplest way of doing this is to introduce a 50% beamsplitter into the signal and idle channels. The method was proposed by Reid and Walls,[50] it was discussed in an earlier paper by Paul[185] and was used in an experiment by, for example, Shih and Alley.[11]

According to (3.3.1), the transformation of the signal beam by the beamsplitter for $\alpha=0$ takes the form $a_{1,2}=(\pm a_{10}+a_{20})/\sqrt{2}$. One of the modes incident on the beamsplitter (a_{20}) is then in the vacuum state. The operators are written here in the Heisenberg representation.

The initial state at the beamsplitter input is described by the vector $|\psi_0\rangle_a=a_{10}^+|00\rangle=|10\rangle_a$. The transformed state can be found by expressing a_{10} in terms of a_1, a_2 with the help of the reverse transformation $a_{10,20}=(a_1\pm a_2)/\sqrt{2}$:

$$|\psi\rangle_a=\frac{1}{\sqrt{2}}(a_1^++a_2^+)|00\rangle=\frac{1}{\sqrt{2}}(|10\rangle_a+|01\rangle_a). \tag{III9}$$

The signal photon has thus become 'smeared out' over the two output modes. The state (III9) is nonfactorizable, but not 'entangled' because it refers to only one photon. We note that the general theory of transformation of quantum fields by different optical devices is discussed in Refs. 33 and 155.

As a result of the above distribution of signal and idle photons, the state vector of the resultant four-mode field takes the form

$$\begin{aligned}|\psi\rangle&=|\psi\rangle_a\otimes|\psi\rangle_b\\&=\tfrac{1}{2}(a_1^++a_2^+)\otimes(b_1^++b_2^+)|0000\rangle\\&=\tfrac{1}{2}(|10\rangle_a|10\rangle_b+|01\rangle_a|01\rangle_b+|10\rangle_a|01\rangle_b\\&\quad+|01\rangle_a|10\rangle_b).\end{aligned} \tag{III10}$$

When it is compared with the required state (3.4.1), the state given by (III10) is unfortunately found to contain 'superfluous' components, i.e., the last two terms, which not only complicate the picture, but also reduces the interference visibility by a factor of 2. The scheme shown in Fig. 1a is in this sense more favorable. The Hamiltonian of (III1) is now replaced with

$$H=i\hbar\Gamma\sum_{k=1}^{2}(a_k^+b_k^+-a_kb_k), \tag{III11}$$

so that instead of (III10) we obtain (3.4.1).

In terms of the compact notation of (3.4.3) and (I2), we can readily verify that the 'entangled' state (3.4.1) has the following properties [see also (I3)–(I5)]:

$$n_1^a|\psi\rangle=n_1^b|\psi\rangle=|10\rangle_a|10\rangle_b/\sqrt{2},$$

$$n_2^a|\psi\rangle=n_2^b|\psi\rangle=|01\rangle_a|01\rangle_b/\sqrt{2}, \tag{III12}$$

$$(n_1^a+n_2^a)|\psi\rangle=(n_1^b+n_2^b)|\psi\rangle=|\psi\rangle,$$

$$n_1^an_2^b|\psi\rangle=0, \tag{III13}$$

$$\langle n_k^a\rangle_\psi=\langle n_k^b\rangle_\psi=\langle n_k^an_k^b\rangle_\psi=1/2, \tag{III14}$$

$$\sigma_-^a\sigma_+^a|\psi\rangle=\sigma_-^b\sigma_+^b|\psi\rangle=\sigma_-^a\sigma_-^b|\psi\rangle=\frac{1}{\sqrt{2}}|01\rangle_a|01\rangle_b, \tag{III15}$$

$$\sigma_+^a\sigma_-^a|\psi\rangle=\sigma_+^b\sigma_-^b|\psi\rangle=\sigma_+^a\sigma_+^b|\psi\rangle=\frac{1}{\sqrt{2}}|10\rangle_a|10\rangle_b, \tag{III16}$$

$$\sigma_z^a|\psi\rangle=\sigma_z^b|\psi\rangle=(|10\rangle_a|10\rangle_b-|01\rangle_a|01\rangle_b)/\sqrt{2}. \tag{III17}$$

We now use the definition (3.4.4) to find the effect of the operators corresponding to the observables A_α, B_β, and $F\equiv A_\alpha B_\beta$:

$$A_\alpha|\psi\rangle=\frac{1}{\sqrt{2}}(|01\rangle_a|10\rangle_be^{i\alpha}+|10\rangle_a|01\rangle_be^{-i\alpha}), \tag{III18}$$

$$F_\varphi|\psi\rangle=\frac{1}{\sqrt{2}}(|01\rangle_a|01\rangle_be^{i\varphi}+|10\rangle_a|10\rangle_be^{-i\varphi}), \tag{III19}$$

$$(AA')^l|\psi\rangle=\frac{1}{\sqrt{2}}[|10\rangle_a|10\rangle_be^{il(\alpha'-\alpha)}+|01\rangle_a|01\rangle_be^{-il(\alpha'-\alpha)}], \tag{III20}$$

$$(FF')^l|\psi\rangle=\frac{1}{\sqrt{2}}[|10\rangle_a|10\rangle_be^{il(\varphi'-\varphi)}+|01\rangle_a|01\rangle_be^{-il(\varphi'-\varphi)}]. \tag{III21}$$

Next, find that the commutators of the observables in the space of states $|\psi\rangle$ are

$$[A,A']=2i\sigma_z^a\sin(\alpha'-\alpha), \tag{III22}$$

$$[F,F']=AA'[B,B']+[A,A']B'B. \tag{III23}$$

According to (3.4.6) or (III20), $A^2=B^2=I$. The square of the Bell operator $S\equiv(AB+A'B+AB'-A'B')/2$ is

$$S^2=I-[A,A'][B,B']/4, \tag{III24}$$

from which we see immediately the effect of the fact that the observables A, A' and B, B' do not commute. Indeed, if we evaluate the average using (III17) and (III22), we obtain the 'quantum' BI[87,89]

$$|\langle S\rangle_\psi|=[1+\langle\sigma_z^a\sigma_z^b\rangle_\psi\sin(\alpha'-\alpha)\times\sin(\beta'-\beta)]^{1/2}\leqslant\sqrt{2}, \tag{III25}$$

which does not contradict experiment.

Next, we must find the joint distribution of the observables A and B, using the standard formula

$$P_{AB}^{\pm\pm}=|\langle\psi|\psi_\alpha^\pm,\psi_\beta^\pm\rangle|^2, \tag{III26}$$

where $|\psi_\alpha^\pm\rangle$ and $|\psi_\beta^\pm\rangle$ are the eigenvalues of A_α and B_β of the form given by (I10), the substitution of which in (III26) gives (3.1.3). The same results follow immediately from definitions such as $P_{AB}^{++}=\langle n_+^an_+^b\rangle\psi$.

Finally, we turn to the quantum description of the parametric amplifier of Gaussian 'seed' noise of intensity N_0 and the influence of this noise on the results of observations obtained with the arrangement of Fig. 1a.

The low-order even moments of the noise field have the form (the odd moments are all zero):

$$\langle a_k^+a_l\rangle_0=\langle b_k^+b_l\rangle_0=N_0\delta_{kl},$$

$$\langle a_k^+b_l^+b_ma_n\rangle_0=N_0^2\delta_{kl}\delta_{mn},$$

$$\langle a_k^+a_l^+a_ma_n\rangle_0=N_0^2(\delta_{km}\delta_{ln}+\delta_{kn}\delta_{lm}), \tag{III27}$$

where k, l, m, $n=1,2$.

The output moments are readily calculated with help of (III5) in the Heisenberg representation:

$$\langle a_k^+a_l\rangle=\langle b_k^+b_l\rangle=N\delta_{kl}, \tag{III28}$$

$$\langle a_kb_l\rangle=\langle b_ka_l\rangle=M\delta_{kl}, \tag{III29}$$

$$\langle a_k^+b_l^+b_ma_n\rangle=N^2\delta_{kn}\delta_{ln}+M^2\delta_{kl}\delta_{mn}, \tag{III30}$$

$$\langle a_k^+a_l^+a_ma_n\rangle=N^2(\delta_{km}\delta_{ln}+\delta_{kn}\delta_{lm}), \tag{III31}$$

where

$$N=u^2N_0+v^2(N_0+1)=(N_0+\tfrac{1}{2})\cosh 2\tau-\tfrac{1}{2}, \tag{III32}$$

$$M=uv(2N_0+1)=(N_0+\tfrac{1}{2})\sinh 2\tau. \tag{III33}$$

We note the following invariant of the transformation (III5):

$$N(N+1)-M^2=N_0(N_0+1)=\text{const}. \tag{III34}$$

Let us now renormalize the observables A_α and B_β so that they assume values in the range $[-1,+1]$ under photon-counting conditions:

$$A_\alpha=(n_+^a-n_-^a)/K,$$

$$B_\beta=(n_+^b-n_-^b)/K, \tag{III35}$$

where

$$K^2=\langle(n_+^a+n_-^a)(n_+^b+n_-^b)\rangle$$

$$=\langle(n_1^a+n_2^a)(n_1^b+n_2^b)\rangle$$

$$=2(2N^2+M^2). \tag{III36}$$

When $N \ll 1$, so that we may take it that, effectively, we have vacuum at the amplifier inputs, we obtain

$$N=\sinh^2\tau, \quad M=N(N+1),$$

$$K^2=2N(3N+1). \tag{III37}$$

Conversely, if $N_0 \gg 1$, the classical result is

$$N=N_0\cosh 2\tau, \quad M=N_0\sinh 2\tau,$$

$$K^2=2N_0^2(1+\cosh^2\tau). \tag{III38}$$

We now return to arbitrary N_0 and form the following operator for a normalized multichannel observable:

$$F\equiv A_\alpha B_\beta=\frac{1}{K^2}\,[\sigma^a_-\sigma^b_- e^{i\varphi}+\sigma^a_-\sigma^b_+ e^{i(\alpha-\beta)}+\text{h.c.}], \tag{III39}$$

with the correlator

$$E\equiv\langle F\rangle=V\cos\varphi. \tag{III40}$$

The interference visibility is given by

$$V=\frac{2M^2}{K^2}=\frac{g-1}{g+1}=\frac{1}{1+(2/\mu^2)} \tag{III41}$$

which is identical to (3.6.3), but was obtained in a different way. In this expression, $g=1+\mu^2$ is the photon bunching parameter and

$$\mu\equiv\frac{M}{N}=\frac{2N_0+1}{N_0\coth\tau+(N_0+1)\tanh\tau}$$

$$\approx\coth\tau \quad \text{for } N_0\ll 1,$$

$$\approx\tanh 2\tau \quad \text{for } N_0\gg 1. \tag{III42}$$

IV. Averaging over random phases in classical models

The evaluation of the integral

$$E_\varphi=\frac{1}{2\pi}\int_{-\pi}^{\pi}\operatorname{sign}[\cos\varphi+\cos(2x)]\,dx$$

$$=\frac{1}{\pi}\int_0^{\pi}\operatorname{sign}(\cos\varphi+\cos x)\,dx \tag{IV1}$$

reduces to the determination of segments on the x axis on which the sign function assumes positive and negative values. Since E_φ is even, it is sufficient to consider the interval $0\leqslant\varphi\leqslant\pi$ on which, within the integration range $0\leqslant x\leqslant\pi$, the integrand is positive for $\cos x>\cos(\pi-\varphi)$, i.e., $0\leqslant x<\pi-\varphi$. The length of a 'negative' interval is φ. By adding together these lengths algebraically, and dividing the result by π, we obtain the final expression given by (3.3.18).

Similarly, evaluation of (4.4.4), i.e.,

$$E_\varphi=\frac{1}{4\pi^2}\iint_{-\pi}^{\pi}F_\varphi\,dx\,dy$$

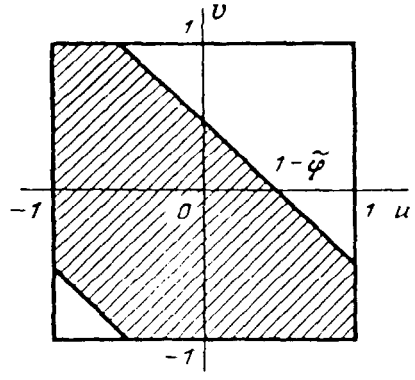

FIG. 9. Evaluation of (4.4.4). The shaded area corresponds to positive arguments of the integrand.

$$=\frac{1}{4\pi^2}\iint_{-\pi}^{\pi}\operatorname{sign}[\cos x\cdot\cos y\cos(\varphi+x+y)]\,dx\,dy \tag{IV2}$$

reduces to the determination of the areas $\Delta_\pm$ of portions $\Lambda_\pm$ of the square $[-\pi,\pi]^2$ on the object plane (x,y) that produce a positive or negative integrand.

By substituting

$$x=\frac{\pi}{2}u, \quad y=\frac{\pi}{2}v, \quad \varphi=\frac{\pi}{2}\tilde\varphi \tag{IV3}$$

we transform the range of integration on the (u,v) plane into the square $[-2,2]^2$ consisting of 16 squares of unit area (cf. Fig. 6). We shall at first by confine our attention to the interval $0\leqslant\tilde\varphi\leqslant 1$.

In the square $u,v\in[0,1]$, the sign of the argument of the sign function in (IV2) is determined by the factor $\cos[(u+v+\tilde\varphi)\pi/2]$ which is positive in the equilateral triangle formed by the u, v axes and the line

$$u+v+\tilde\varphi=1. \tag{IV4}$$

The length of a side of this triangle (Fig. 9) is $1-\tilde\varphi$, so that the contribution of the square to E_φ is

$$\Delta_+-\Delta_-=2\Delta_+-1, \quad \Delta_+=(1-\tilde\varphi)^2/2. \tag{IV5}$$

An equal but opposite contribution is provided by the square u, $v\in[-1,0]$. The two thus cancel out. The total number of such pairs that transform into each other by the rule $(u,v)\to(-u,-v)$ and are arranged chessbard fashion is eight (Fig. 6).

A representative of the other family of the remaining eight squares that do not mutually cancel out is $u\in[0,1]$, $v\in[-1,1]$. In this, a 'negative triangle' is defined by

$$u+v+\tilde\varphi\in[1,2]. \tag{IV6}$$

The area of this triangle is $\Delta_-=\tilde\varphi^2/2$, so that the contribution of the square is $1-\tilde\varphi^2$. Hence, recalling the normalization condition, we have

$$E_\varphi=(1-\tilde\varphi^2)/2, \quad |\tilde\varphi|\leqslant 1. \tag{IV7}$$

Similarly, we can calculate E_φ for $|\tilde{\varphi}|\in[1,2]$, which follows from (IV7) after replacing $|\tilde{\varphi}|$ with $2-|\tilde{\varphi}|$ and changing the sign. In (IV7), and henceforth, we replace $\tilde{\varphi}$ with $|\tilde{\varphi}|$ because E_φ is even.

To conclude, we evaluate the N-tuple integral in (5.4.8), taking into account the parametric relation

$$\sum_{n=1}^{N}\lambda_n=0$$

and the random jumps by π in (5.4.9).

Since all the λ_n can only be equal to 0 or π, we have $\sin\lambda_n=0$ and

$$\cos(\alpha_N+\lambda_n)=\cos\left(\alpha_N-\sum_{n=1}^{N-1}\lambda_n\right)$$
$$=\cos\alpha_N\prod_{n=1}^{N-1}\cos\lambda_n. \qquad \text{(IV8)}$$

Consequently, the integrand can be factorized and we obtain the product of $N-1$ integrals of the form

$$\frac{1}{2}\int_{-\pi}^{\pi}\operatorname{sign}[\cos\lambda_n\cdot\cos(\alpha_n+\lambda_n)][\delta(\lambda_n)+\delta(\lambda_n-\pi+0)]d\lambda_n$$
$$=\operatorname{sign}\cos\alpha_n, \qquad \text{(IV9)}$$

from which there follows the final expression given by (5.4.10).

[1]A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935) [Russ. transl., Usp. Fiz. Nauk **16**, 440 (1936)].
[2]D. Bohm, *Quantum Theory*, Prentice-Hall, Englewood Cliffs, 1951 [Russ. transl., Fizmatgiz, Moscow, 1961].
[3]J. S. Bell, Physics **1**, 195 (1964).
[4]J. F. Clauser and A. Shimony, Rep. Prog. Phys. **41**, 1881 (1978).
[5]A. A. Grib, Usp. Fiz. Nauk **142**, 619 (1984) [Sov. Phys. Usp. **27**, 284 (1984)].
[6]A. I. Akhiezer and R. V. Polovin, Usp. Fiz. Nauk **107**, 463 (1984) [Sov. Phys. Usp. **15**, 500 (1972)].
[7]V. P. Demutskiĭ and R. V. Polovin, Usp. Fiz. Nauk **162**, 95 (1992) [Sov. Phys. Usp. **35**, 857 (1992)].
[8]D. Home and M. A. B. Whitaker, Phys. Rep. **210**, 223 (1992).
[9]D. Home and F. Selleri, Nuovo Cimento **14**, 1 (1991).
[10]A. Sudbery, *Quantum Mechanics and the Particles of Nature*, Cambridge University Press, Cambridge, 1986 [Russ. transl., Mir, Moscow, 1989].
[11]C. O. Alley and Y. H. Shih, Proc. 2nd Internat. Symp. on Foundations of Quantum Mechanics, edited by M. Namuki *et al.*, Tokyo, 1987; Phys. Rev. Lett. **61**, 2921 (1988).
[12]Z. Y. Ou and L. Mandel, Phys. Rev. Lett. **61**, 50 (1988).
[13]J. G. Rarity and P. B. Tapster, Phys. Rev. Lett. **64**, 2495 (1990).
[14]D. M. Greenberger, M. A. Horne, and A. Zeilinger, *Bell's Theorem: Quantum Theory and Conceptions of the Universe*, edited by M. Kufatos, Kluwer, Dordrecht, 1989, p. 69.
[15]D. M. Greenberger, M. A. Horne, A. Shimony, and A. Zeilinger, Am. J. Phys. **58**, 1131 (1990).
[16]N. D. Mermin, Phys. Rev. Lett. **65**, 1838 (1990).
[17]S. M. Roy and V. Singh, Phys. Rev. Lett. **67**, 2761 (1991).
[18]M. Ardehali, Phys. Rev. A **46**, 5375 (1992).
[19]A. V. Belinsky and D. N. Klyshko, Phys. Lett. A, in press (1993).
[20]L. Hardy, Phys. Lett. A **160**, 1 (1991).
[21]D. N. Klyshko, Phys. Lett. A **172**, 399 (1993).
[22]L. Hardy, Phys. Lett. A **167**, 17 (1992).
[23]H. J. Bernstein, D. M. Greenberger, M. A. Horne, and A. Zeilinger, **47**, 78 (1993).
[24]A. Peres, Phys. Lett. A **151**, 107 (1990); **163**, 243 (1992).
[25]N. D. Mermin, Phys. Rev. Lett. **65**, 3373 (1990).
[26]S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967).
[27]H. R. Brown and G. Svetlichny, Found. Phys. **20**, 1379 (1990).
[28]A. Elby, *ibid.*, **20**, 1389 (1990).
[29]A. Stairs, Phil. Sci. **50**, 578 (1983).
[30]S. J. Freedman and J. F. Clauser, Phys. Rev. Lett. **28**, 938 (1972).
[31]A. Aspect, J. Dalibar, and G. Roger, Phys. Rev. Lett. **49**, 1804 (1982).
[32]W. Perrie, A. J. Duncan, H. J. Beyer, and H. Kleinpoppen, Phys. Rev. Lett. **54**, 1790 (1985).
[33]A. V. Belinsky and D. N. Klyshko, Laser Phys. **2**, 112 (1992).
[34]D. N. Klyshko, Phys. Lett. A **163**, 349 (1992).
[35]A. V. Belinsky and D. N. Klyshko, Phys. Lett. A **166**, 303 (1992).
[36]A. Shimony, Sci. Am. **258**, 2 (1988) [Russ. transl. in *V Mire Nauki* No. 3, 22 (1988)].
[37]B. d'Espagnat, Sci. Am. **241**, 128 (1979).
[38]G. E. Stedman, Am. J. Phys. **53**, 1143 (1985).
[39]D. N. Klyshko, Phys. Lett. A **132**, 299 (1988).
[40]J. F. Clauser, M. A. Horne, A. Shomony, R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[41]J. F. Clauser and M. A. Horne, Phys. Rev. D **10**, 526 (1974).
[42]H. R. Brown and R. Q. Twiss, Nature **177**, 27 (1956).
[43]R. J. Glauber, Phys. Rev. **131**, 2766 (1963); Les Houches and Summer School, Gordon and Breach, N.Y., 1964
[44]E. Santos, Phys. Rev. Lett. **66**, 1388 (1991); **68**, 2700 (1992).
[45]E. Santos, Phys. Rev. A **46**, 3646 (1992).
[46]M. A. Horne, A. Shimony, and A. Zeilinger, **62**, 2209 (1989).
[47]M. A. Horne and A. Zeilinger, Proc. Symp. on Foundations of Modern Physics, edited by P. Lachti and P. Mittelstaedt, World Scientific, Singapore, 1985, p. 435.
[48]P. D. Drummond, Phys. Rev. Lett. **50**, 1407 (1983).
[49]M. D. Reid and D. F. Walls, Phys. Rev. Lett. **53**, 955 (1984).
[50]M. D. Reid and D. F. Walls, Phys. Rev. A **34**, 1260 (1986).
[51]M. D. Reid and P. D. Drummond, Phys. Rev. Lett. **60**, 2731 (1988).
[52]M. Zukowski and J. Pykacz, Phys. Lett. A **127**, 1 (1988).
[53]Z. Y. Ou, C. K. Hong, and L. Mandel, Opt. Commun. **67**, 159 (1988).
[54]Z. Y. Ou, S. F. Pereira, H. J. Kimble, and K. C. Peng, Phys. Rev. Lett. **68**, 3663 (1992).
[55]P. Grangier, M. J. Potasek, and B. Yurke, Phys. Rev. A **38**, 3132 (1988).
[56]M. D. Reid and D. Drummond, Phys. Rev. Lett. **60**, 2447 (1988).
[57]M. D. Reid, Phys. Rev. A **40**, 913 (1989).
[58]B. J. Oliver and C. R. Stroud, Phys. Rev. A **135**, 407 (1989).
[59]S. M. Tan, M. D. Holland, and D. F. Walls, Opt. Commun. **77**, 285 (1990).
[60]S. M. Tan and D. F. Walls, Opt. Commun. **71**, 235 (1989).
[61]M. S. Chubarov and E. P. Nikolaev, Phys. Lett. A **110**, 199 (1985).
[62]Z. Y. Ou, C. K. Hong, and L. Mandel, Phys. Lett. A **122**, 11 (1987).
[63]Z. Y. Ou, Phys. Rev. A **37**, 1607 (1988).
[64]D. N. Klyshko, Usp. Fiz. Nauk **154**, 133 (1988) [Sov. Phys. Usp. **31**, 74 (1988)].
[65]D. N. Klyshko, Usp. Fiz. Nauk **158**, 327 (1989) [Sov. Phys. Usp. **32**, 355 (1989)].
[66]D. N. Klyshko, Phys. Lett. A, **128**, 133 (1988).
[67]L. M. Davis, Phys. Lett. A **140**, 275 (1989).
[68]G. E. Stedman, J. Mod. Opt. **36**, 1461 (1989).
[69]M. Zukowski, Phys. Lett. A **150**, 136 (1990).
[70]M. Zukowski and A. Zeilinger, Phys. Lett. A **155**, 69 (1991).
[71]B. J. Oliver and C. R. Stroud, J. Opt. Soc. Am. B **4**, 1426 (1987).
[72]A. Joshi and S. V. Lawande, Phys. Rev. A **44**, 716 (191).
[73]H. P. Stapp, Found. Phys. **18**, 427 (1988).
[74]B. d'Espagnat, Phys. Rep. **110**, 201 (1984).
[75]L. Hardy, Phys. Lett. A **161**, 21 (1991).
[76]A. C. Elitzur, S. Popescu, and D. Rohrlich, Phys. Lett. A **162**, 25 (1992).
[77]P. H. Eberhard, Nuovo Cimento B **38**, 75 (1977).
[78]S. J. Feingold and A. Peres, J. Phys. A **13**, 3187 (1980).
[79]H. P. Stapp, Phys. Rev. Lett. **49**, 1470 (1982).
[80]H. P. Stapp, Phys. Rev. Lett. **53**, 306 (1985).
[81]P. H. Eberhard, Phys. Rev. Lett. **49**, 1474 (1982).
[82]A. Fine, Phys. Rev. Lett. **48**, 291 (1982).

[83]A. Fine, Phys. Rev. Lett. **49**, 242 (1982).
[84]S. N. Bismas, S. R. Choudhury, and K. Datta, Phys. Lett. A **153**, 279 (1991).
[85]R. K. Clifton, J. N. Butterfield, and M. L. J. Redhead, J. Phil. Sci. **41**, 5 (1990).
[86]M. A. B. Whitaker and J. G. Dennison, Phys. Lett. A **133**, 466 (1988).
[87]B. S. Cirelson, Lett. Math. Phys. **4**, 93 (1980).
[88]L. J. Landau, Phys. Lett. A **120**, 54 (1987).
[89]M. E. Andas, Phys. Lett. A **167**, 6 (1992).
[90]S. L. Braunstein, A. Mann, and M. Revzen, Phys. Rev. Lett. **68**, 3259 (1992).
[91]S. Popescu and D. Rohrlich, Phys. Lett. A **169**, 411 (1992).
[92]S. J. Summers and R. F. Werner, J. Math. Phys. **28**, 24480 (1987).
[93]A. O. Barut and P. Meystre, Phys. Lett. A **105**, 458 (1984).
[94]R. Livi, Lett. Nuovo Cimento **19**, 189 (1977).
[95]N. D. Mermin, Phys. Rev. D **22**, 356 (1980).
[96]N. D. Mermin and G. M. Schwarz, Found. Phys. **12**, 101 (1982).
[97]M. Ogren, Phys. Rev. D **27**, 1766 (1983).
[98]A. Garg and N. D. Mermin, Phys. Rev. Lett. **49**, 901 (1982).
[99]A. Garg and N. D. Mermin, Phys. Rev. Lett. **49**, 1294 (1982).
[100]A. Garg and N. D. Mermin, Phys. Rev. D **27**, 339 (1983).
[101]M. Ardehali, Phys. Rev. D **444**, 3336 (1991).
[102]A. Peres, Phys. Rev. A **46**, 4413 (1992).
[103]N. Gisin, Phys. Lett. A **154**, 201 (1991).
[104]N. Gisin and A. Peres, Phys. Lett. A **162**, 15 (1992).
[105]S. Popescu and D. Rohrlich, Phys. Lett. A **162**, 293 (1992).
[106]S. M. Barnett and S. J. D, Phoenix, Phys. Lett. A **167**, 233 (1992).
[107]S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **61**, 662 (1988).
[108]S. L. Braunstein and C. M. Caves, Ann. Phys. **202**, 22 (1990).
[109]S. M. Barnett and S. J. D. Phoenix, Phys. Rev. A **44**, 535 (1991).
[110]B. W. Schumacker, Phys. Rev. A **44**, 7047 (1991).
[111]L. Hardy, Phys. Rev. Lett. **68**, 2981 (1992).
[112]L. Hardy and E. J. Squires, **168**, 169 (1992).
[113]C. Pagonis and R. Clifton, Phys. Lett. A **168**, 100 (1992).
[114]H. Clifton and P. Niemann, Phys. Lett. A **166**, 177 (1992).
[115]C. J. Dewdney, J. Phys. A **25**, 3615 (1992).
[116]C. Pagonis, M. L. G. Redhead, and R. Clifton, Phys. Rev. A **155**, 441 (1991).
[117]P. Heywood and M. L. G. Redhead, Found. Phys. **13**, 481 (1983).
[118]N. D. Mermin, Am. J. Phys. **58**, 731 (1990).
[119]N. D. Mermin, Phys. Today **43**, 9 (1990).
[120]H. S. Choi, Phys. Lett. A **153**, 285 (1991).
[121]B. Yurke, Phys. Rev. A **46**, 2229 (1992).
[122]M. D. Reid and W. J. Munro, Phys. Rev. Lett. **69**, 997 (1992).
[123]M. Zukowski, Phys. Lett. A **157**, 198 (1991)
[124]M. Zukowski, Phys. Lett. A **157**, 203 (1991)
[125]I. Pitowski, Phys. Lett. A **156**, 137 (1991)
[126]I. Pitowski, Phys. Lett. A **166**, 292 (1992)
[127]F. Herbut, Phys. Lett. A **163**, 5 (1992).
[128]A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[129]C. H. Bennet, G. Brassard, and N. D. Mermin, Phys. Rev. Lett. **68**, 557 (1992).
[130]C. H. Bennet, Phys. Rev. Lett. **68**, 3121 (1992).
[131]A. K. Ekert, J. G. Rarity, P. R. Tapster, and G. M. Palma, Phys. Rev. Lett. **69**, 1293 (1992).
[132]L. Mandel, J. Opt. Soc. Am. B **1**, 108 (1984).
[133]C. K. Hong, S. R. Friberg, and L. Mandel, Appl. Opt. **24**, 3877 (1985).
[134]S. F. Seward, P. R. Tapster, J. G. Walker, and J. G. Rarity, Quant. Opt. **3**, 201 (1991).
[135]J. G. Rarity, P. R. Tapster, J. G. Walker, and S. F. Seward, Appl. Opt. **29**, 2939 (1990).
[136]D. N. Klyshko and A. N. Penin, Usp. Fiz. Nauk **152**, 653 (1987) [Sov. Phys. Usp. **30**, 716 (1987)].
[137]G. Kh. Kutaeva and A. N. Penin, Kvant. Elektron. **19**, 816 (1992) [Sov. J. Quantum Electron. **22**, 756 (1992)].
[138]A. N. Penin and A. V. Sergienko, Appl. Opt. **30**, 3583 (1991).
[139]G. C. Ghirardi, A. Rimini, and T. Weber, Lett. Nuovo Cimento **27**, 293 (1980).
[140]P. J. Bussey, Phys. Lett. A **90**, 9 (1982).
[141]D. Dieks, Phys. Lett. A **92**, 271 (1982).
[142]L. E. Ballentine and J. P. Jarret, Am. J. Phys. **55**, 696 (1987).
[143]G. C. Ghirardi, R. Grassi, A. Rimini, and T. Weber, Europhys. Lett. **6**, 95 (1988).
[144]J. Polchinski, Phys. Rev. Lett. **66**, 397 (1991).
[145]N. Gisin, Phys. Lett. A **143**, 1 (1990).
[146]A. Valentini, Phys. Rev. A **42**, 639 (1990).
[147]W. M. Muynck, Phys. Lett. A **114**, 65 (1986).
[148]W. M. Muynck and Phys. Lett. A **142**, 187 (1989).
[149]W. M. Muynck and O. Abu-Zeid, Phys. Lett. A **100**, 193 (1984).
[150]A. Berthelot, Nuovo Cimento **57**, 193 (1980).
[151]P. Ratall, Found. Phys. **13**, 555 (1983).
[152]W. Muckenheim, Lett. Nuovo Cimento **35**, 300 (1982).
[153]K. Wodkiewicz, Phys. Lett. A **129**, 1 (1988); G. S. Agarwal, D. Home, and W. Schleich, Phys. Lett. A **170**, 359 (1992).
[154]D. N. Klyshko, Phys. Lett. A **146**, 93 (1990).
[155]D. N. Klyshko, Phys. Lett. A **137**, 334 (1989).
[156]D. N. Klyshko, Laser Phys. **2**, 997 (1992).
[157]A. V. Belinskiĭ, Zh. Eksp. Teor. Fiz. **103**, 1538 (1993).
[158]W. H. Louisell, *Radiation and Noise in Quantum Electronics*, McGraw-Hill, N.Y. 1964 [Russ. transl. Nauka, Moscow, 1972].
[159]P. Curruthers and M. Nieto, Rev. Mod. Phys. **40**, 411 (1968).
[160]C. Su and K. Wodkiewicz, **44**, 6097 (1991).
[161]C. T. Lee, Phys. Rev. A **44**, 2775 (1991).
[162]D. C. Burnham and D. L. Weinberg, Phys. Rev. Lett. **25**, 84 (1970).
[163]B. Ya. Zel'dovich and D. N. Klyshko, Pis'ma Zh. Eksp. Teor. Fiz. **9**, 69 (1969) [JETP Lett. **9**, 40 (1969)].
[164]J. D. Franson, Phys. Rev. Lett. **62**, 2205 (1989).
[165]Z. Y. Ou, X. Y. Zou, L. G. Wang, and L. Mandel, Phys. Rev. Lett. **65**, 321 (1990).
[166]Z. Y. Ou and L. Mandel, J. Opt. Soc. Am. B **7**, 2127 (1990).
[167]S. A. Akhmanov, Usp. Fiz. Nauk **149**, 361 (1986) [Sov. Phys. Usp. **29**, 589 (1986)].
[168]S. A. Akhmanov, A. V. Belinskiĭ, and A. S. Chirkin, Kvant. Elektron. **15**, 873 (1988) [Sov. J. Quantum Electron. **18**, 560 (1988)].
[169]S. A. Akhmanov, A. V. Belinskiĭ, and A. S. Chirkin, Opt. Spektrosk. **66**, 738 (1989) [Opt. Spectrosc. (USSR) **66**, 432 (1989)].
[170]S. A. Akhmanov, A. V. Belinskiĭ, and A. S. Chirkin, *New Physical Principles for Optical Data Processing* [in Russian], edited by S. A. Akhmanov and M. A. Vorontsov, Nauka, Moscow, 1990, p. 83–194.
[171]S. A. Akmanov, Izv. Vuz, Ser. Radiofiz. **4**, 769 (1961).
[172]S. A. Akhmanov and V. P. Komolov, Vestn. Mosk. Univ., Ser. **3**, Fiz. Astron. **7**, 95 (1966).
[173]S. A. Akhmanov and R. V. Khokhlov, Radiotekh. Elektron. **6**, 1813 (1961).
[174]Yu. E. D'yakov, Radiotekh. Elektron. **8**, 1812 (1963).
[175]Yu. E. D'yakov, Radiotekh. Elektron. **8**, 2034 (1963).
[176]S. A. Akhmanov, Yu. D. D'yakov, and A. S. Chirkin, *Introduction to Statistical Radiophysics and Optics* [in Russian], Nauka, Moscow, 1981.
[177]A. V. Belinskiĭ and D. N. Klyshko, Zh. Eksp. Teor. Fiz. **102**, 1116 (1992) [Sov. Phys. JETP **75**, 606 (1992)].
[178]D. N. Klyshko, Phys. Lett. A **146**, 471 (1990).
[179]P. V. Elyutin and D. N. Klyshko, Phys. Lett. A **149**, 241 (1990).
[180]W. Vogel and W. Schleich, Phys. Rev. A **44**, 7642 (1991).
[181]J. W. Noh, A. Fougeres, and L. Mandel, Phys. Rev. Lett. **67**, 1426 (1991).
[182]J. W. Noh, A. Fougeres, and L. Mandel, Phys. Rev. A **45**, 424 (1992).
[183]M. Freyberger and W. Schleich, Phys. Rev. A **47**, 30 (1993).
[184]V. I. Smirnov, *A Course of Higher Mathematics* [in Russian], GITTL, 1957, Vol. 3, p. 248.
[185]H. Paul, Opt. Acta **1**, 28 (1981).
[186]J. Brendel, E. Mohler, and W. Martienssen, Europhys. Lett. **20**, 575 (1992).
[187]Y. H. Shi, A. V. Sergienko, and M. H. Rubin, Phys. Rev. A **47**, 1288 (1993).
[188]H. P. Stapp, Phys. Rev. A **47**, 847 (1993).
[189]W. H. Zurek, S. Habib, and J. P. Paz, Phys. Rev. Lett. **70**, 1187 (1993).
[190]J. F. Williams, M. Kumar, and A. T. Stelbovics, Phys. Rev. Lett. **70**, 1240 (1993).
[191]D. N. Smithey, M. Beck, M. G. Raymer, and A. Faridani, Phys. Rev. Lett. **70**, 1244 (1993).
[192]A. Elby and and M. P. Jonts, Phys. Lett. A **171**, 11 (1992).
[193]A. Mann, M. Revzen, and W. Schleich, Phys. Rev. A **46**, 5363 (1992).
[194]F. Selleri, *Quantum Paradoxes and Physical Reality*, Kluwer, Dordrecht, 1990.
[195]D. Morgan, Sci. Am. **267**, 1 (1992) [Russ. transl. in *V mire nauki* No. 9–10, 70 (1992)].

[196]M. Belsley, D. T. Smithey, and M. G. Raymer, Phys. Rev. A **46**, 414 (1992).
[197]T. S. Larchuk, R. A. Campos, J. G. Rarity, P. R. Tapster, E. Jakeman, B. E. A. Saleh, and M. C. Teich, Phys. Rev. Lett. **70**, 1603 (1993).
[198]M. Vinduska, Phys. Lett. A, to be published (1993).
[199]M. Ardehali, Phys. Rev. A **47**, 1633 (1993).
[200]H. Scherer and P. Busch, Phys. Rev. A **47**, 1647 (1993).
[201]S. L. Braunstein and A. Mann, Phys. Rev. A **47**, 2427 (1993).
[202]P. G. Kwiat, A. M. Steinberg, and R. Y. Chiao, Phys. Rev. A **47**, 2472 (1993).
[203]C. H. Bennett, G. Brassard, C. Crepeau, R. Jorsa, A. Peres, and W. K. Wooters, Phys. Rev. Lett. **70**, 1895 (1993).

Translated by S. Chomet