

Two interferometric complementarities

Gregg Jaeger

Department of Physics, Boston University, Boston, Massachusetts 02215

Abner Shimony

Departments of Physics and Philosophy, Boston University, Boston, Massachusetts 02215

Lev Vaidman

Department of Physics and Astronomy, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel

(Received 1 August 1994)

Two interferometric complementarities are formulated and demonstrated, and relations between them are analyzed. The first relates the distinguishability $D(\mathcal{P})$ of the path of propagation of a particle (where \mathcal{P} is the preparation of an ensemble) to the fringe visibility v_1 when amplitudes from two paths are combined. It is shown that $[\max D(\mathcal{P})]^2 + v_1^2 = 1$, where the maximum is taken over all preparations compatible with a fixed density operator. The second complementarity relates the visibility of one-particle interference fringes to the visibility v_{12} of two-particle fringes: $v_{12}^2 + v_1^2 \leq 1$. With a suitable extension of operations on the pair of particles, this inequality is strengthened to an equality.

PACS number(s): 03.65.Bz, 07.60.Ly

I. INTRODUCTION

The first of the two interferometric complementarities that we shall study in this paper has been roughly known since the early days of quantum mechanics: it is the complementarity of the *distinguishability* of the path of a particle and the *visibility* of interference fringes formed when amplitudes from two paths are combined with variable phase factors. In a famous paper, Bohr [1] analyzed two variants of the double-slit experiment. In one variant, the diaphragm pierced with two slits is free to move so that the slit through which the particle passes can be determined by measuring the recoil of the diaphragm. In the other, the diaphragm is fixed, so that the path cannot be determined. But only in the latter arrangement is an interference pattern exhibited. He concluded, “we are presented with a choice of *either* tracing the path of a particle *or* observing interference effects.”

Wootters and Zurek [2], followed by Bartell [3], initiated the study of arrangements intermediate between the two extremes considered by Bohr. They found that measurements can be made to determine retrodictively with high probability which slit each particle of the ensemble traversed, without completely obliterating the interference pattern formed by the impingement of the particles on a screen, and they found a complementarity between the quantity of “partial information” (suitably defined) and the visibility of the interference pattern. Bartell remarked, however, that “Although such a ratio suggests a fairly high probability of a trajectory through slit A , it cannot, of course, be rigidly interpreted in terms of individual photons going through slit A four times out of five.”

This conceptual difficulty suggests the value of posing a somewhat different though related problem that concerns prediction rather than retrodiction. Suppose that an en-

semble of particles given by a preparation \mathcal{P} emerges from a diaphragm into two beams A and A' , as shown in Fig. 1. One could insert either of two kinds of apparatus downstream from the slit. One [Fig. 1(a)] is an apparatus for observing the interference between beams A and A' , brought together on a viewing screen S . If this type of apparatus is chosen, the interesting physical quantity is the *visibility* v_1 of the interference pattern caused by the preparation \mathcal{P} . The other kind of apparatus [Fig. 1(b)] consists of ideal detectors inserted in beams A and A' , and the interesting quantity is the *distinguishability* of the path, which can be defined in the following way. Consider any measurements that can be made consistently with the preparation \mathcal{P} , for example, by examining auxiliary systems correlated with the particles of the ensemble; formulate a strategy for predicting, for each particle in the

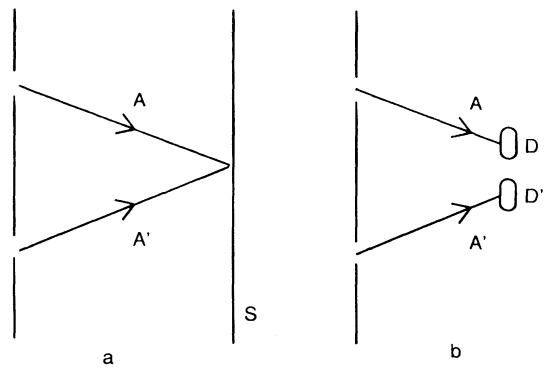


FIG. 1. Two apparatuses for observing particle beams A and A' emerging from a diaphragm having two slits: (a) A and A' are brought together on a viewing screen S and (b) A and A' are made to impinge on detectors D and D' placed in their paths.

ensemble, whether it will be detected in beam A or in beam A' , where the strategy can of course make use of knowledge of the preparation \mathcal{P} in addition to the results of the measurements bearing on the particle of interest. These predictions can be checked by the ideal detectors in the beams A and A' . The optimum strategy, given \mathcal{P} , is the one for which the probability of a correct prediction has the maximum value p_{\max} , which clearly is also the strategy that makes the probability of an incorrect prediction have a minimum value q_{\min} . We propose as the measure of path distinguishability, given preparation \mathcal{P} ,

$$D(\mathcal{P}) = p_{\max} - q_{\min} = 2p_{\max} - 1. \quad (1)$$

Any monotonic function of $D(\mathcal{P})$ would be an acceptable candidate for a definition of path distinguishability, but ours has the obvious virtue of possessing the interval $0-1$ as its range and it has less obvious virtues, which will appear later, of yielding simple relations to visibility and to the density matrix associated with the preparation.

A paper of Greenberger and YaSin [4] addressed the probability of prediction rather than retrodiction [in the discussion preceding their Eq. (7)] and they proposed a concept of path distinguishability that is the same as ours, but in a restricted range of preparations. They only considered preparations in which no observations on auxiliary systems can yield information useful for predicting the path of a particle of interest and therefore the strategy of prediction must be based entirely on the preparation. We call such preparations “simple.” The experimental situation that they studied consisted of an incident neutron beam split at the first slab of a neutron interferometer into two beams A and A' , which are recombined at a point downstream. Their analysis applies, however, to any preparation of an idealized ensemble of particles, each of which we shall refer to as “particle 1,” propagating in beams A and/or A' , where the locution “and/or” refers in a condensed way to the possibility of quantum-mechanical superposition. Their idealization consists in assuming that the Hilbert space H_1 associated with each particle 1 is two dimensional, spanned by vectors $|A\rangle$ and $|A'\rangle$, where the former completely characterizes propagation in beam A and the latter completely characterizes propagation in beam A' . If the ensemble is a pure simple case, with all particles in the same quantum state $|\psi\rangle$,

$$|\psi\rangle = c|A\rangle + c'|A'\rangle, \quad (2)$$

then the path distinguishability (using our notation, not theirs) is

$$D = |(|c|^2 - |c'|^2)|. \quad (3)$$

If instead the ensemble is a mixture of pure simple cases, which they describe as “partially coherent,” there is a more complicated expression for D [given by their Eq. (15)]. (They do not consider the statistical simple case in which a density operator ρ is given on H_1 rather than definite proportions of pure states $|\psi_i\rangle$.) In either the pure or the mixed simple case, there is a complementarity between path distinguishability and visibility. In the pure

simple case, the complementarity can be expressed by an equality

$$D^2 + v_1^2 = 1; \quad (4a)$$

in the mixed simple case it is expressed by the inequality

$$D^2 + v_1^2 \leq 1. \quad (4b)$$

Since entering a definite beam is a particlelike property, while exhibiting an interference pattern is a wavelike property, Greenberger and YaSin use the notations P and W , respectively, for our D and v_1 and they interpret (4a) and (4b) as expressing wave-particle complementarity.

In our Sec. II we extend the analysis of Greenberger and YaSin by considering a larger class of preparations. We add to the pure and mixed simple cases those cases defined by density operators, which we call “statistical simple cases.” Furthermore, quantum mechanics postulates that the Hilbert space representing states of a composite system $1+2$ is the direct product of the Hilbert spaces H_1 and H_2 associated with 1 and 2 separately. If $|\Theta\rangle \in H_1 \otimes H_2$ cannot be factored into a product of a vector in H_1 and a vector in H_2 , then the state represented by $|\Theta\rangle$ is entangled. One can consider an ensemble of pairs $1+2$, all characterized by the same $|\Theta\rangle$, and then focus attention only upon the particle 1 from each pair. There is a standard procedure [5] for “tracing out” the variables associated with particle 2 in order to derive from $|\Theta\rangle$ a density operator for the ensemble of particles 1. We call the ensemble of particles 1 a “pure entangled case” (rather than the usual locution “improper mixture” [6]). An ensemble formed by taking several pure entangled cases, each defined by a $|\Theta_k\rangle$ in $H_1 \otimes H_2$ with respective proportions w_k , will be called a “mixed entangled case” and an ensemble defined by giving a density operator σ on $H_1 \otimes H_2$ but no further information will be called a “statistical entangled case.” We show in Sec. II that the complementarity in the strong form of Eq. (4a) holds if D is taken to be the distinguishability $D(|\Theta\rangle)$ calculated in a pure entangled case determined by $|\Theta\rangle$ and the inequality (4b) holds if D is calculated in the mixed or statistical entangled cases.

We also make a connection with a recent paper of Mandel [7], who defines a quantity P_D , which is his proposed measure of path distinguishability, in terms of the density operator ρ on H_1 . We show that there are distinct preparations \mathcal{P} and \mathcal{P}' , both determining the same density operator ρ , such that

$$D(\mathcal{P}) \neq D(\mathcal{P}'), \quad (5a)$$

$$D(\mathcal{P}) > P_D. \quad (5b)$$

Accordingly we assert that the natural measure of path distinguishability should be a function of preparations rather than of density operators. We also show in Sec. II that if $\max D(\mathcal{P})$ is the maximum distinguishability for all preparations \mathcal{P} determining a fixed density operator ρ and $|\Theta\rangle$ is any vector in $H_1 \otimes H_2$ that yields this ρ for the particle 1 ensemble, then

$$\max D(\mathcal{P}) = D(|\Theta\rangle). \quad (6)$$

The complementarity expressed in Eqs. (4a) and (4b) can be rewritten with greater generality as

$$[\max D(\mathcal{P})]^2 + v_1^2 = 1 \quad (7a)$$

and

$$[D(\mathcal{P})]^2 + v_1^2 \leq 1. \quad (7b)$$

In Sec. III we shall study a second complementarity, between one-particle interferometric visibility v_1 and two-particle interferometric visibility v_{12} . The experimental arrangement in which this complementarity is exhibited is shown schematically in Fig. 2. Particle 1 is in beams A and/or A' and particle 2 is in beams B and/or B' , and each pair is in a pure quantum state

$$\begin{aligned} |\Theta\rangle = & \gamma_1 |A\rangle|B\rangle + \gamma_2 |A\rangle|B'\rangle \\ & + \gamma_3 |A'\rangle|B\rangle + \gamma_4 |A'\rangle|B'\rangle, \end{aligned} \quad (8a)$$

where

$$|\gamma_1|^2 + |\gamma_2|^2 + |\gamma_3|^2 + |\gamma_4|^2 = 1, \quad (8b)$$

$|A\rangle$ and $|A'\rangle$ are vectors in H_1 representing states of propagation in the beams A and A' , and $|B\rangle$ and $|B'\rangle$ are analogous vectors in H_2 .

Beams A and A' are brought together into a passive lossless transducer T_1 , which feeds two output beams U_1 and L_1 , while beams B and B' are brought together into another passive lossless transducer T_2 , which feeds output beams U_2 and L_2 , as indicated in Fig. 2. The output beams are equipped with ideal detectors. Both T_1 and T_2 can be varied, and one can study the probability $P(U_1 U_2)$ of joint detection in beams U_1 and U_2 and the analogous probabilities $P(U_1 L_2)$, $P(U_2 L_1)$, and $P(L_1 L_2)$, as well as the probabilities of single detection $P(U_i)$ and $P(L_i)$ ($i=1,2$). Interferometry is concerned with the variation of these probabilities as the transducers T_1 and T_2 are varied. We are particularly interested in the one-particle fringe visibility

$$V_i = \frac{[P(U_i)]_{\max} - [P(U_i)]_{\min}}{[P(U_i)]_{\max} + [P(U_i)]_{\min}} \quad (i=1,2) \quad (9)$$

and in the two-photon fringe visibility V_{12} , which is analogous to V_i , but with some important differences which will be made explicit later. The word "fringe," of course, is borrowed from classical optical interferometry, where it is descriptive, but we use the term generically to

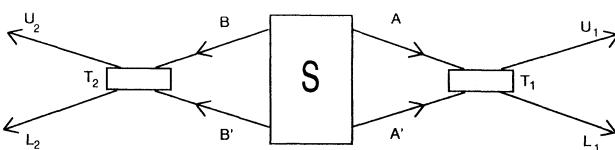


FIG. 2. Schematic two-particle interferometer. Two pairs of beams A, A' and B, B' impinge on passive lossless transducers T_1 and T_2 , respectively, and emerge in beams U_1, L_1 and U_2, L_2 .

refer to variability of probability as the transducers are varied.

The remarkable phenomena of two-particle interferometry result from the fact that when $|\Theta\rangle$ is entangled, it can happen that

$$P(U_1 U_2) \neq P(U_1)P(U_2), \quad (10)$$

and likewise for the other joint probabilities $P(U_1 L_2)$, $P(U_2 L_1)$, and $P(L_1 L_2)$. Because the transducers T_1 and T_2 operate independently, the correlated behavior of 1 and 2 manifested in the inequality (10) is a striking instance of quantum nonlocality. Although two-particle interferometry is only about a decade old, it has been intensively studied theoretically and experimentally [8].

It was noticed by Horne and Zeilinger [9] that when the two-particle visibility is unity, the one-particle visibility is zero, and conversely. A systematic investigation of intermediate cases was carried out by Jaeger, Horne, and Shimony [10], who demonstrated that in a large family of states $|\Theta\rangle$ a complementarity holds between one-particle and two-particle visibility:

$$v_{12}^2 + v_1^2 \leq 1. \quad (11a)$$

(The significance of the small letter v , as contrasted with the capital V of Eq. (9), will be explained in Sec. III.) In this paper, we prove that the inequality (11a) holds for all $|\Theta\rangle$ and in addition that a stronger complementarity, in the form of an equality, is valid:

$$V_{12}^2 + V_1^2 = 1. \quad (11b)$$

The two complementarities—between path distinguishability and single-particle visibility and between one-particle and two-particle visibilities—are intimately connected. The more entangled $|\Psi\rangle$ is, the stricter is the bound on the one-particle visibility. The reason, roughly, is that the phase information between the basis vectors $|A\rangle$ and $|A'\rangle$ of H_1 is carried by correlated vectors of H_2 and therefore observations made only upon particle 1 of each pair 1+2 cannot fully extract this phase information. On the other hand, a high degree of entanglement has two consequences: it entails high two-particle fringe visibility and it permits good inferences about the path of particle 1 to be made on the basis of measurements of particle 2. It is not surprising, then, that the one-particle visibility should enter in the same way in both of the complementarity relations (7a) and (7b) and (11a) and (11b), which on the surface seem remote from each other. We note, finally, that these complementarities are derived without any approximations.

II. PATH DISTINGUISHABILITY

In Sec. I a general definition was proposed for the path distinguishability $D(\mathcal{P})$ in a certain class of experimental arrangements, as a function of the preparation \mathcal{P} . A general expression was given for $D(\mathcal{P})$ in Eq. (1). We also gave the following classification of preparations: (I) simple cases, subdivided into (Ia) pure simple cases, (Ib) mixed simple cases, and (Ic) statistical simple cases; and (II) entangled cases, subdivided into (IIa) pure entangled

cases, (IIb) mixed entangled cases, and (IIc) statistical entangled cases. We know of no preparation that is statistically complete, in the sense of providing sufficient information to calculate the expectation values of all quantum-mechanical observables of the system of interest, that does not fall into one of these six cases.

Each of these preparations determines a density operator ρ unequivocally. For our purposes, it is most useful to write the density matrix of ρ in the $|A\rangle, |A'\rangle$ basis

$$\begin{aligned}\rho_{11} &= \langle A | \rho | A \rangle, \quad \rho_{12} = \langle A | \rho | A' \rangle, \\ \rho_{21} &= \langle A' | \rho | A \rangle, \quad \rho_{22} = \langle A' | \rho | A' \rangle.\end{aligned}\quad (12)$$

The density operator ρ satisfies the standard conditions of self-adjointness, positivity, and having trace 1, which imply that

$$\rho_{11} \geq 0, \quad (13a)$$

$$\rho_{22} \geq 0, \quad (13b)$$

$$\rho_{11} + \rho_{22} = 1, \quad (13c)$$

$$\rho_{21} = \rho_{12}^*, \quad (13d)$$

$$|\rho_{12}| \leq (\rho_{11}\rho_{22})^{1/2}. \quad (13e)$$

It will be useful to have an explicit expression for ρ_{ij} in each of the cases Ia, Ib, Ic, IIa, IIb, and IIc.

Case Ia (pure simple). There is a pure quantum state of each member of the ensemble represented by $|\psi\rangle$ of Eq. (2). Then

$$\rho_{11} = |c|^2, \quad \rho_{12} = cc'^*, \quad \rho_{21} = c^*c', \quad \rho_{22} = |c'|^2. \quad (14)$$

Case Ib (mixed simple). There is a mixture of subensembles, each in the pure simple case described by $|\psi_k\rangle$,

$$|\psi_k\rangle = c_k |A\rangle + c'_k |A'\rangle, \quad (15)$$

and the proportions are w_k (summing to unity). Then

$$\begin{aligned}\rho_{11} &= \sum_k w_k |c_k|^2, \quad \rho_{12} = \sum_k w_k c_k c'_k, \\ \rho_{21} &= \sum_k w_k c_k^* c'_k, \quad \rho_{22} = \sum_i w_k |c'_k|^2.\end{aligned}\quad (16)$$

Case Ic (statistical simple). The ensemble is defined by the statement of ρ .

Case IIa (pure entangled). There is a pure state $|\Theta\rangle$ of 1+2, which can always be written as

$$|\Theta\rangle = N^{1/2} |A\rangle |\chi\rangle + \lambda N^{1/2} |A'\rangle |\chi'\rangle + N'^{1/2} |A'\rangle |\chi'\rangle, \quad (17a)$$

where $|\chi\rangle$ and $|\chi'\rangle$ are orthonormal vectors of H_2 (spanning a two-dimensional subspace that we shall call \bar{H}_2), N and N' are non-negative,

$$N(1 + |\lambda|^2) + N' = 1, \quad (17b)$$

and

$$\lambda = |\lambda| e^{i\theta_\lambda}. \quad (17c)$$

Without loss of generality we can assume

$$|\lambda| \leq 1 \quad (17d)$$

(an assumption that may require interchanging the labels A and A'). The method of “tracing out” [5] yields

$$\begin{aligned}\rho_{11} &= N, \\ \rho_{12} &= \rho_{21}^* = N |\lambda| e^{-i\theta_\lambda}, \\ \rho_{22} &= N |\lambda|^2 + N' = 1 - N.\end{aligned}\quad (18)$$

Case IIb (mixed entangled). There is a mixture of pure states of 1+2, represented by normalized (but not necessarily orthonormal) vectors $|\Theta_k\rangle \in H_1 \otimes H_2$, each of the form

$$\begin{aligned}|\Theta_k\rangle &= N_k^{1/2} |A_k\rangle |\chi_k\rangle + \lambda_k N_k^{1/2} |A'_k\rangle |\chi'_k\rangle \\ &\quad + N_k'^{1/2} |A'_k\rangle |\chi'_k\rangle,\end{aligned}\quad (19a)$$

where A_k is A or A' and A'_k is A' or A , the choice being made so as to guarantee the analog of (17d)

$$|\lambda_k| \leq 1. \quad (19b)$$

The analogs of (17b) and (17c), with the subscript k placed upon N , N' , λ , and Θ_λ , are also assumed. The proportions of the pure states $|\Theta_k\rangle$ in the mixture are w_k . If $A_k = A$ and $A'_k = A'$, then $\rho_{ij}^{(k)}$ is the analog of Eq. (18):

$$\begin{aligned}\rho_{11}^{(k)} &= N_k, \\ \rho_{12}^{(k)} &= \rho_{21}^{(k)*} = N_k |\lambda_k| e^{-i\theta_{k\lambda}}, \\ \rho_{22}^{(k)} &= N_k |\lambda_k|^2 + N'_k = 1 - N_k.\end{aligned}\quad (20a)$$

However, if $A_k = A'$ and $A'_k = A$, then

$$\begin{aligned}\rho_{11}^{(k)} &= N_k |\lambda_k|^2 + N'_k = 1 - N_k, \\ \rho_{21}^{(k)} &= \rho_{12}^{(k)*} = N_k |\lambda_k| e^{-i\theta_{k\lambda}}, \\ \rho_{22}^{(k)} &= N_k.\end{aligned}\quad (20b)$$

Then the density matrix ρ_{ij} for the mixture (uniformly associating index 1 with A and index 2 with A') is

$$\rho_{ij} = \sum_k w_k \rho_{ij}^{(k)}. \quad (21)$$

Case IIc (statistical entangled). A density operator σ is given in $H_1 \otimes H_2$. It is always possible to express σ in the diagonal form

$$\sigma = \sum_k w_k |\Theta_k\rangle \langle \Theta_k|, \quad (22)$$

where the $|\Theta_k\rangle$ are orthonormal vectors in $H_1 \otimes H_2$, $w_k > 0$, and the w_k sum to unity. (Note that orthonormality is assumed in this case, but was not in case IIb.) The tracing out procedure yields a statistical matrix ρ_{ij} for particle 1, which is exactly the same as Eq. (21), where $\rho_{ij}^{(k)}$ are defined as in Eqs. (20a) and (20b).

We now proceed to calculate $D(\mathcal{P})$ for \mathcal{P} belonging to each of the cases Ia, Ib, Ic, IIa, IIb, and IIc. Only case IIa is difficult.

In case Ia we shall designate \mathcal{P} by the $|\psi\rangle$ with which it is associated. Because the case is simple, the strategy

for prediction can only depend upon $|\psi\rangle$ and there are only two nonrandomized strategies: (i) predict A in each case and (ii) predict A' in each case. Obviously strategy (i) is optimum if $|c| > |c'|$, strategy (ii) is optimum if $|c'| > |c|$, and the choice is indifferent if $|c| = |c'|$. A randomized strategy, sometimes using (i) and sometimes using (ii), is never an improvement [11]. Hence

$$D(|\psi\rangle) = |(|c|^2 - |c'|^2)| , \quad (23)$$

as Greenberger and YaSin already found [4] and we reported in Eq. (3) above. Equation (23) can be rewritten usefully by noting that Eq. (14) implies

$$(|c|^2 - |c'|^2)^2 = 1 - 4|\rho_{12}|^2 . \quad (24)$$

Hence

$$D(|\psi\rangle) = (1 - 4|\rho_{12}|^2)^{1/2} . \quad (25)$$

In case Ib, clearly the optimum strategy in the mixed ensemble is to use the optimum strategy for the k th subensemble, when it is known that the particle of interest belongs to the k th subensemble. If we designate \mathcal{P} in the mixed simple case by $\{|\psi_k\rangle, w_k\}$, then

$$\begin{aligned} D(\{|\psi_k\rangle, w_k\}) &= \sum_k w_k |(|c_k|^2 - |c'_k|^2)| \\ &= \sum_k w_k (1 - 4|\rho_{12}^{(k)}|^2)^{1/2} . \end{aligned} \quad (26)$$

We note that Eq. (26) is more general than Eq. (24) of Greenberger and YaSin.

We represent the preparation \mathcal{P} in the statistical simple case Ic by the same letter ρ that designates the density operator. Since no information about the ensemble other than ρ is provided, one cannot subdivide the ensemble into subensembles for which different strategies are optimum. The only nonrandomized strategies are those labeled (i) and (ii) in case Ia. The probability of detection in A is $\text{Tr}(\rho P_A)$, where $P_A = |A\rangle\langle A|$, and the probability of detection in A' is $\text{Tr}(\rho P_{A'})$, where $P_{A'} = |A'\rangle\langle A'|$. Whether strategy (i) or (ii) is followed one obtains

$$\begin{aligned} D(\rho) &= p_{\max} - q_{\min} = |\text{Tr}(\rho P_A) - \text{Tr}(\rho P_{A'})| \\ &= |\rho_{11} - \rho_{22}| = |2\rho_{11} - 1| . \end{aligned} \quad (27)$$

We designate the preparation \mathcal{P} in the pure entangled case IIa by $|\Theta\rangle$. By a lengthy argument that we defer,

$$D(|\Theta\rangle) = (1 - 4|\rho_{12}|^2)^{1/2} , \quad (28)$$

where ρ_{12} is the matrix element given in Eq. (18). We point out, however, two remarkable features of Eq. (28). The first is that $D(|\Theta\rangle)$ is independent of the choice of $|\Theta\rangle$ within the class of vectors of $H_1 \otimes H_2$ that yield the same $|\rho_{12}|$. The second (a corollary of the first) is an independence from the detailed character of the Hilbert space H_2 and hence of the system 2 entangled with particle 1.

In case IIb we designate the preparation \mathcal{P} by $\{|\Theta_k\rangle, w_k\}$. By the same remark made for case Ib we write

$$D(\{|\Theta_k\rangle, w_k\}) = \sum_k w_k (1 - 4|\rho_{12}^{(k)}|^2)^{1/2} . \quad (29)$$

We represent the preparation \mathcal{P} in the statistical entangled case IIc by the same letter σ that designates the density operator describing the ensemble of composite systems 1+2. As in case Ic, we cannot divide the ensemble into subensembles that call for different strategies of prediction. Since the probability of detection in A is $\text{Tr}[\sigma(P_A \otimes 1)]$ and of detection in A' is $\text{Tr}[\sigma(P_{A'} \otimes 1)]$, we have

$$\begin{aligned} D(\sigma) &= |\text{Tr}[\sigma(P_A \otimes 1)] - \text{Tr}[\sigma(P_{A'} \otimes 1)]| \\ &= |\text{Tr}(\rho P_A) - \text{Tr}(\rho P_{A'})| = |\rho_{11} - \rho_{22}| = |2\rho_{11} - 1| , \end{aligned} \quad (30)$$

where ρ is obtained from σ by tracing out.

We now fill the gap in our exposition by demonstrating Eq. (28), which states the distinguishability in the pure entangled case IIa. The most general strategy is to take advantage of the correlations between 1 and 2 implicit in $|\Theta\rangle$ and use the result of measuring a bivalent operator on H_2 . With no loss of generality, the bivalent operator can be restricted to be a projection on H_2 and the strategy is the following: when projection E is measured on particle 2, predict that its partner, particle 1, will be detected in A or A' according to whether the result of the measurement is 1 or 0. If this rule is followed, then the probability of a correct prediction minus the probability of a false prediction is

$$\begin{aligned} D_E(|\Theta\rangle) &= \langle \Theta | P_A \otimes E | \Theta \rangle + \langle \Theta | P_{A'} \otimes (1-E) | \Theta \rangle \\ &\quad - [\langle \Theta | P_A \otimes (1-E) | \Theta \rangle + \langle \Theta | P_{A'} \otimes E | \Theta \rangle] \\ &= \langle \Theta | (P_{A'} - P_A) \otimes 1 | \Theta \rangle \\ &\quad + 2 \langle \Theta | (P_A - P_{A'}) \otimes E | \Theta \rangle . \end{aligned} \quad (31)$$

By Eqs. (17a) and (17b), the first term on the right-hand side of Eq. (31) is

$$\langle \Theta | (P_{A'} - P_A) \otimes 1 | \Theta \rangle = |\lambda|^2 N + N' - N = 1 - 2N . \quad (32)$$

The effect of E operating on the vectors $|\chi\rangle$ and $|\chi'\rangle$ of Eq. (17a) is the following:

$$E|\chi\rangle = c|\chi\rangle + c'|\chi'\rangle + c''|\chi''\rangle , \quad (33a)$$

$$E|\chi'\rangle = d|\chi\rangle + d'|\chi'\rangle + d''|\bar{\chi}'\rangle , \quad (33b)$$

where $|\chi''\rangle$ and $|\bar{\chi}'\rangle$ are unit vectors orthogonal to both $|\chi\rangle$ and $|\chi'\rangle$ (i.e., they lie outside the subspace \bar{H}_2). Since E is a projection

$$\langle \chi | E | \chi \rangle = \langle \chi | E^2 | \chi \rangle = \langle \chi | E | E \chi \rangle , \quad (34)$$

and hence

$$c = |c|^2 + |c'|^2 + |c''|^2 ; \quad (35a)$$

likewise

$$d' = |d|^2 + |d'|^2 + |d''|^2 . \quad (35b)$$

Furthermore, c and d' are real and

$$c' = d^* = |c'| e^{i\theta_{c'}} . \quad (35c)$$

Then

$$\begin{aligned}\langle \Theta | (P_A - P_{A'}) \otimes E | \Theta \rangle &= N(1 - |\lambda|^2)c - \lambda^* N^{1/2} N'^{1/2} d - \lambda N^{1/2} N'^{1/2} c' - N'd' \\ &= N(1 - |\lambda|^2)c - 2|\lambda||c'|N^{1/2}N'^{1/2}\cos(\theta_\lambda + \theta_{c'}) - N'd' .\end{aligned}\quad (36)$$

The phase angles that maximize this expression and therefore maximize $D_E(|\Theta\rangle)$ satisfy $\cos(\theta_\lambda + \theta_{c'}) = -1$, so that

$$\begin{aligned}\max \langle \Theta | (P_A - P_{A'}) \otimes E | \Theta \rangle \\ = N(1 - |\lambda|^2)c + 2|\lambda||c'|N^{1/2}N'^{1/2} - N'd' .\end{aligned}\quad (37)$$

The maximum solution of Eq. (35a) and the minimum solution of Eq. (35b) are

$$c = \frac{1}{2}\{1 + [1 - 4(|c'|^2 + |c''|^2)]^{1/2}\} , \quad (38a)$$

$$d' = \frac{1}{2}\{1 - [1 - 4(|c'|^2 + |d''|^2)]^{1/2}\} . \quad (38b)$$

For a given $|c'|$ these are respectively maximized and minimized by taking $|c''| = |d''| = 0$, which shows that the projection operator E of the optimum strategy remains a projection when it is restricted to the two-dimensional subspace \bar{H}_2 . Because of Eqs. (38a) and (38b), the maximization in Eq. (37) is obtained by varying a single real parameter

$$k = |c'|^2 \quad (39)$$

in the expression

$$\begin{aligned}\max \langle \Theta | (P_A - P_{A'}) \otimes E | \Theta \rangle \\ = N(1 - |\lambda|^2)\frac{1}{2}[1 + (1 - 4k)^{1/2}] + 2|\lambda|N^{1/2}N'^{1/2}k^{1/2} \\ - N'\frac{1}{2}[1 - (1 - 4k)^{1/2}] .\end{aligned}\quad (40)$$

Setting the derivative with respect to k to zero yields

$$(1 - 4k)^{1/2} = K[N(1 - |\lambda|^2) + N'] , \quad (41a)$$

$$k^{1/2} = K|\lambda|N^{1/2}N'^{1/2} , \quad (41b)$$

where the common factor K is fixed by Eq. (17b) to be

$$K = (1 - 4N^2|\lambda|^2)^{-1/2} . \quad (41c)$$

$$\begin{aligned}D(|\Theta\rangle) = \max D_E(|\Theta\rangle) &= \max \{[\langle \Theta | P_A \otimes |\phi\rangle \langle \phi | \Theta \rangle - \langle \Theta | P_{A'} \otimes |\phi\rangle \langle \phi | \Theta \rangle] \\ &+ [\langle \Theta | P_{A'} \otimes |\phi'\rangle \langle \phi' | \Theta \rangle - \langle \Theta | P_A \otimes |\phi'\rangle \langle \phi' | \Theta \rangle]\} \\ &= [\langle \psi_1 | P_A | \psi_1 \rangle - \langle \psi_1 | P_{A'} | \psi_1 \rangle] + [\langle \psi_2 | P_{A'} | \psi_2 \rangle - \langle \psi_2 | P_A | \psi_2 \rangle] = \sum_{i=1}^2 w_i D(|\psi_i\rangle) ,\end{aligned}\quad (46)$$

where the maximizing $|\phi\rangle, |\phi'\rangle$ are used to define

$$|\psi_1\rangle = \frac{\langle \phi | \Theta \rangle}{\| \langle \phi | \Theta \rangle \|} \in H_1 , \quad (47a)$$

$$|\psi_2\rangle = \frac{\langle \phi' | \Theta \rangle}{\| \langle \phi' | \Theta \rangle \|} \in H_2 , \quad (47b)$$

$$w_1 = \|\langle \phi | \Theta \rangle\|^2 , \quad (48a)$$

$$w_2 = \|\langle \phi' | \Theta \rangle\|^2 . \quad (48b)$$

From Eqs. (31), (32), (40), (41a)–(41c), and (18) one obtains

$$\begin{aligned}D(|\Theta\rangle) = \max D_E(|\Theta\rangle) &= (1 - 4N^2|\lambda|^2)^{1/2} \\ &= (1 - 4|\rho_{12}|^2)^{1/2} ,\end{aligned}\quad (42)$$

completing the derivation of Eq. (28).

It is interesting to note that for any pure entangled preparation $|\Theta\rangle$ there is a mixed simple case $\{|\psi_i\rangle, w_i\}$ such that

$$D(\{|\psi_i\rangle, w_i\}) = D(|\Theta\rangle) . \quad (43)$$

To demonstrate this fact, we recall that the projection operator E maximizing $D_E(|\Theta\rangle)$ was shown [see remark after Eqs. (38a) and (38b)] to have the property that its restriction to \bar{H}_2 , denoted by \bar{E} , is a projection on \bar{H}_2 . In other words, either

$$\bar{E} = 1 \text{ (the identity operator)} , \quad (44a)$$

$$\bar{E} = 0 \text{ (the null operator)} , \quad (44b)$$

or

$$\bar{E} = |\phi\rangle \langle \phi| , \quad (44c)$$

where $|\phi\rangle$ is a normalized vector in \bar{H}_2 . But the use of a strategy based on Eq. (44a) could be optimum only if the coefficient $N'^{1/2}$ in Eq. (17a) is zero, in which case $|\Theta\rangle$ is a product state and particle 1 is in a pure state, and hence this instance of case IIa reduces effectively to a pure simple case Ia. Likewise, if a strategy based on Eq. (44b) is optimum, then the coefficient $N^{1/2}$ must be zero and again $|\Theta\rangle$ is a product state. If Eq. (44c) holds, then there is a normalized vector $|\phi'\rangle \in \bar{H}_2$ that is orthogonal to $|\phi\rangle$ and

$$1 - E = |\phi'\rangle \langle \phi'| . \quad (45)$$

Hence, by Eq. (31)

These vectors $|\psi_i\rangle$ and weights w_i satisfy Eq. (43), as promised. There is nothing surprising in this agreement, because the result of measuring E is to achieve a reduction of the wave packet $|\Theta\rangle$ for each composite system 1+2 of the ensemble, yielding two subensembles. In the subensemble corresponding to $E = +1$, each particle 1 is in the state described by $|\psi_1\rangle$, and in the subensemble corresponding to $E = -1$, each particle is in $|\psi_2\rangle$.

We now are in a position to compare our measure of

distinguishability $D(\mathcal{P})$ with Mandel's P_D . Mandel points out that a density operator ρ on a two-dimensional Hilbert space can be uniquely expressed in the form

$$\rho = P_{\text{ID}}\rho_{\text{ID}} + P_D\rho_D, \quad (49)$$

where ρ_D is diagonal in the $|A\rangle, |A'\rangle$ basis and ρ_{ID} is the density operator of a pure quantum state. The subscript D denotes "distinguishable" and ID denotes "indistinguishable." Specifically, if ρ_{ij} is the ij matrix element of ρ in the $|A\rangle, |A'\rangle$ basis, then

$$\rho_D = \begin{bmatrix} \rho_{11} & 0 \\ 0 & \rho_{22} \end{bmatrix}, \quad (50a)$$

$$\rho_{\text{ID}} = \begin{bmatrix} \rho_{11} & \rho_{12}/P_{\text{ID}} \\ \rho_{21}/P_{\text{ID}} & \rho_{22} \end{bmatrix}, \quad (50b)$$

$$P_{\text{ID}} = \frac{|\rho_{12}|}{(\rho_{11}\rho_{22})^{1/2}}, \quad (50c)$$

$$P_D = 1 - \frac{|\rho_{12}|}{(\rho_{11}\rho_{22})^{1/2}}, \quad (50d)$$

and a vector that yields ρ_{ID} is

$$|\psi\rangle = \rho_{11}^{1/2}|A\rangle + e^{-i\sigma}\rho_{22}^{1/2}|A'\rangle, \quad (50e)$$

where

$$\rho_{12} = |\rho_{12}|e^{i\sigma}. \quad (50f)$$

Mandel proposes P_D as his measure of path distinguishability in the arrangement of Fig 1. As a rationale for this definition he notes that there is an experimental scheme that generates a proportion ρ_{11} of particles in A and a proportion ρ_{22} in A' , and the ensemble so generated would be described by ρ_D .

Now consider the preparation \mathcal{P} that consists of three subensembles: proportion $\rho_{11}P_D$ in $|A\rangle$, proportion $\rho_{22}P_D$ in $|A'\rangle$, and proportion P_{ID} in the $|\psi\rangle$ of Eq. (50e). By following Mandel's decomposition we see that the preparation \mathcal{P} determines the original density operator ρ . But using the $D(\mathcal{P})$ of case Ib above we find

$$\begin{aligned} D(\mathcal{P}) &= \rho_{11}P_D \times 1 + \rho_{22}P_D \times 1 + P_{\text{ID}}(|\rho_{11} - \rho_{22}|) \\ &= P_D + P_{\text{ID}}(|\rho_{11} - \rho_{22}|). \end{aligned} \quad (51)$$

Clearly, we obtain inequality (5b), unless either P_{ID} is zero or ρ_{11} equals ρ_{22} . We claim that our measure of distinguishability $D(\mathcal{P})$ is preferable to Mandel's P_D because ours takes into account the difference between the coefficients of $|A\rangle$ and $|A'\rangle$ in the superposition $|\psi\rangle$ as a ground for making predictions about detection in A or A' .

It is interesting to exhibit another preparation \mathcal{P}' of class Ib with the same ρ but different distinguishability.

Assume that $\rho_{11} > \rho_{22}$ and let \mathcal{P}' be the following mixture: proportion $(\rho_{11} - \rho_{22})P_D$ in $|A\rangle$, proportion $\frac{1}{2}\rho_{22}P_D$ in $(1/\sqrt{2})(|A\rangle + |A'\rangle)$, proportion $\frac{1}{2}\rho_{22}P_D$ in $(1/\sqrt{2})(|A\rangle - |A'\rangle)$, and proportion P_{ID} in $|\psi\rangle$. Then

$$\begin{aligned} D(\mathcal{P}') &= (\rho_{11} - \rho_{22})P_D \times 1 + 0 + 0 + P_{\text{ID}}(\rho_{11} - \rho_{22}) \\ &< D(\mathcal{P}) \end{aligned} \quad (52)$$

(unless P_D or P_{ID} is zero). We have thus proved the statement made in Sec. I that $D(\mathcal{P})$ depends essentially upon the preparation \mathcal{P} , not just upon ρ .

We now demonstrate Eq. (6), stated in Sec. I, asserting that for a given ρ , the maximum distinguishability is given by the pure entangled case IIa, i.e., by a $|\Theta\rangle \in H_1 \otimes H_2$, that yields ρ (of course, as just demonstrated, there is a mixed simple case Ib with the same distinguishability). We proceed by two lemmas.

Lemma 1. If $\rho = w\rho^{(1)} + (1-w)\rho^{(2)}$, with $0 < w < 1$, and $\rho, \rho^{(1)}, \rho^{(2)}$ are all two-dimensional density matrices, then

$$\begin{aligned} (1-4|\rho_{12}|^2)^{1/2} &\geq w(1-4|\rho_{12}^{(1)}|^2)^{1/2} \\ &\quad + (1-w)(1-4|\rho_{12}^{(2)}|^2)^{1/2}. \end{aligned} \quad (53)$$

Proof. $|\rho_{12}|^2 = w^2x^2 + (1-w)^2y^2 + 2w(1-w)xy\cos\theta$, where $x = |\rho_{12}^{(1)}|$, $y = |\rho_{12}^{(2)}|$, and $\theta = \arg\rho_{12}^{(2)} - \arg\rho_{12}^{(1)}$. By the general properties of the density matrix, the real numbers $|\rho_{12}|$, x , and y all lie in the interval $[0, \frac{1}{2}]$. Then

$$\begin{aligned} 1 &\geq 4xy + (1-4x^2)^{1/2}(1-4y^2)^{1/2} \\ &\geq 4xy\cos\theta + (1-4x^2)^{1/2}(1-4y^2)^{1/2}, \end{aligned}$$

whence

$$\begin{aligned} |\rho_{12}|^2 &\leq \frac{1}{4} - \frac{1}{4}w^2(1-4x^2) - \frac{1}{4}(1-w^2)(1-4y^2) \\ &\quad - \frac{1}{2}w(1-w)(1-4x^2)^{1/2}(1-4y^2)^{1/2}, \end{aligned}$$

so that

$$(1-4|\rho_{12}|^2)^{1/2} \geq w(1-4x^2)^{1/2} + (1-w)(1-4y^2)^{1/2}.$$

Lemma 2. If $\rho = \sum_{i=1}^{\infty} w_i \rho^{(i)}$, where ρ and all $\rho^{(i)}$ are two-dimensional density matrices, $0 < w_i < 1$, and $\sum w_i = 1$, then

$$(1-4|\rho_{12}|^2)^{1/2} \geq \sum_{i=1}^{\infty} w_i(1-4|\rho_{12}^{(i)}|^2)^{1/2}. \quad (54)$$

Proof by induction. As the induction hypothesis, suppose the inequality (54) is valid whenever the summation has only n terms. Consider now

$$\rho = \sum_{i=1}^{n+1} w_i \rho^{(i)}, \quad \sum_{i=1}^{n+1} w_i = 1.$$

Then

$$\begin{aligned} \sum_{i=1}^{n+1} w_i(1-4|\rho_{12}^{(i)}|^2)^{1/2} &= \kappa \sum_{i=1}^n \frac{w_i}{\kappa}(1-4|\rho_{12}^{(i)}|^2)^{1/2} + w_{n+1}(1-4|\rho_{12}^{(n+1)}|^2)^{1/2} \\ &\leq \kappa \left[\sum_{i=1}^n \frac{w_i}{\kappa} \right] (1-4|\rho_{12}^{(n+1)}|^2)^{1/2} + w_{n+1}(1-4|\rho_{12}^{(n+1)}|^2)^{1/2} \end{aligned}$$

by using the induction hypothesis, where

$$\kappa = \sum_{i=1}^n w_i = 1 - w_{n+1}, \quad \rho' = \sum_{i=1}^n \frac{w_i}{\kappa} \rho^{(i)}.$$

By lemma 1

$$\begin{aligned} \kappa(1-4|\rho'_{12}|^2)^{1/2} + w_{n+1}(1-4|\rho_{12}^{(n+1)}|^2)^{1/2} \\ \leq (1-4|\kappa\rho'_{12} + w_{n+1}\rho_{12}^{(n+1)}|^2)^{1/2} \\ = (1-4|\rho_{12}|^2)^{1/2}. \end{aligned}$$

Hence the induction is complete and the inequality (54) holds when the summation is finite, i.e., when

$$\sum_{i=1}^N w_i = 1, \quad \sum_{i=1}^N w_i \rho^{(i)} = \rho$$

for arbitrary N . By continuity the inequality is established for an infinite summation, proving lemma 2.

To complete the proof of the inequality (6), we note that if \mathcal{P} is a mixed simple case Ib, i.e., a mixture of $|\psi_k\rangle$ with proportion w_k , then the density operator ρ determined by \mathcal{P} is the weighted sum of the density operators $\rho^{(k)}$ determined by $|\psi_k\rangle$,

$$\rho = \sum_k w_k \rho^{(k)}.$$

By Eqs. (54) (lemma 2) and (26),

$$\begin{aligned} D(\{|\psi_k\rangle, w_k\}) &= \sum_k w_k (1-4|\rho_{12}^{(k)}|^2)^{1/2} \\ &\leq (1-4|\rho_{12}|^2)^{1/2}. \end{aligned} \quad (55)$$

Likewise, if \mathcal{P} is a mixed entangled case, i.e., mixtures of $|\Theta_k\rangle$ with proportions w_k , then again the ρ determined by \mathcal{P} is the weighted sum of $\rho^{(k)}$ determined by $|\Theta\rangle$, and Eqs. (54) and (26) imply

$$D(\{|\Theta_k\rangle, w_k\}) \leq [1-4|\rho_{12}|^2]^{1/2}. \quad (56)$$

In the statistical simple case Ic

$$D(\rho) = |\rho_{11} - \rho_{22}|.$$

By Eq. (13e)

$$\begin{aligned} D(\rho) &= [\rho_{11}^2 - 2\rho_{11}\rho_{22} + \rho_{22}^2]^{1/2} \\ &= [\rho_{11}^2 + 2\rho_{11}\rho_{22} + \rho_{22}^2 - 4\rho_{11}\rho_{22}]^{1/2} \\ &= [1-4\rho_{11}\rho_{22}]^{1/2} \\ &\leq [1-4|\rho_{12}|^2]^{1/2}. \end{aligned} \quad (57)$$

In the statistical simple entangled case IIc we proved

$$D(\sigma) = |\rho_{11} - \rho_{22}|,$$

where ρ_{ij} is the density matrix for particle 1 derived from σ by tracing out. The argument of Eq. (57) then shows

$$D(\sigma) \leq [1-4|\rho_{12}|^2]^{1/2}. \quad (58)$$

In the pure simple case Ia and the pure entangled case IIa

$$D(\mathcal{P}) = [1-4|\rho_{12}|^2]^{1/2},$$

by Eqs. (25) and (28). We have therefore surveyed all preparations \mathcal{P} compatible with a specified ρ , and the inequality (6) holds.

As a simple corollary, we have an inequality governing Mandel's P_D :

$$P_D \leq (1-4|\rho_{12}|^2)^{1/2}, \quad (59)$$

because we have already exhibited in the inequality (5b) that P_D is bounded from above by $D(\mathcal{P})$, where \mathcal{P} is a preparation of class Ib. It is also straightforward to check the inequality (59) directly.

Finally, we turn to the complementarity of path distinguishability and fringe visibility. The latter is evaluated by Mandel in Ref. [7] in terms of the density matrix

$$v_1 = 2|\rho_{12}|. \quad (60)$$

To make the present exposition self-contained, we shall give a brief derivation of Eq. (60). Figure 3 adds to the arrangement of Fig. 1 some mirrors that direct beams A and A' to an ideal symmetric beam splitter, with transmittivity and reflectivity equal to $1/\sqrt{2}$, the output from the beam splitter being beams U and L , and it inserts a phase shift τ into beam A . Then

$$|A\rangle \rightarrow e^{i\tau} \frac{1}{\sqrt{2}} [i|U\rangle + |L\rangle], \quad (61a)$$

$$|A'\rangle \rightarrow \frac{1}{\sqrt{2}} [|U\rangle + i|L\rangle]. \quad (61b)$$

It is easy to check that if

$$|\phi(\tau)\rangle = \frac{1}{\sqrt{2}} [e^{-i(\tau+\pi/2)}|A\rangle + |A'\rangle], \quad (62)$$

then a particle initially in $|\phi(\tau)\rangle$ will pass with certainty into U . Hence the proposition "passage into U due to the symmetric beam splitter with phase shift τ " is naturally represented by the projection operator

$$\begin{aligned} Q(\tau) &= |\phi(\tau)\rangle \langle \phi(\tau)| \\ &= \frac{1}{2} |A\rangle \langle A| + \frac{1}{2} e^{-i(\tau+\pi/2)} |A\rangle \langle A'| \\ &\quad + \frac{1}{2} e^{i(\tau+\pi/2)} |A'\rangle \langle A| + \frac{1}{2} |A'\rangle \langle A'|. \end{aligned} \quad (63)$$

Given the density operator ρ for an ensemble of particles and also the operator $Q(\tau)$, we can express the probability of passage into beam U as

$$P(\tau) = \text{Tr}[Q(\tau)\rho] = \frac{1}{2}\rho_{11} + \frac{1}{2}\rho_{22} - |\rho_{12}|\sin(\tau + \sigma), \quad (64)$$

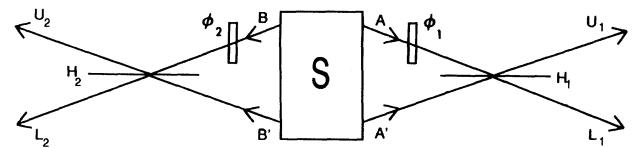


FIG. 3. Schematic two-particle interferometer. Beam pairs A, A' and B, B' emerge from source S , the upper beams A and B pass through variable phase shifters ϕ_1 and ϕ_2 , respectively, and the pairs of beams impinge on lossless beam splitters H_1 and H_2 , respectively, emerging in beams U_1, L_1 and U_2, L_2 .

where ρ_{12} is as defined in (50f). In the standard manner, the visibility v_1 is defined as

$$v_1 = \frac{[P(U_1)]_{\max} - [P(U_1)]_{\min}}{[P(U_1)]_{\max} + [P(U_1)]_{\min}}, \quad (65)$$

where the maximum and minimum are computed as τ is varied. The obvious result is Eq. (60).

The conjunction of Eq. (6), which states the maximum path distinguishability for a specified ρ , with Eqs. (28) and (60) yields Eq. (7a), as promised in Sec. I. The inequality (7b) is an immediate consequence.

III. INTERFEROMETRY

A schematic arrangement for two-particle interferometry was described and depicted in Fig. 2. We now give a mathematical formulation of the arrangement of Fig. 2. Roughly, a transducer is passive if no particle exits from it that has not entered and it is lossless if an incoming particle is certain to exit. We shall bypass the problems of analyzing these concepts by assuming that a passive lossless transducer is represented by a unitary unimodular mapping, the domain of which is the space of input states and the counterdomain of which is the space of output states. In the case of T_1 , the domain is the subspace spanned by vectors $|A\rangle$ and $|A'\rangle$, which represent (though not uniquely) propagation in the beams A and A' , and the counterdomain is the subspace spanned by the vectors $|U_1\rangle$ and $|L_1\rangle$, which represent (though not uniquely) propagation in beams U_1 and L_1 . In the case of T_2 the domain is spanned by $|B\rangle$ and $|B'\rangle$ and the counterdomain by $|U_2\rangle$ and $|L_2\rangle$, which have analogous interpretations. No confusion will result from using T_i ($i=1,2$) equivocally to denote the transducer and the associated unitary mapping. The most general state of the composite system $1+2$, given that photon 1 is in beams A and/or A' and photon 2 is in B and/or beam B' , is the symmetrized version of

$$|\Theta\rangle = \gamma_1|A\rangle|B\rangle + \gamma_2|A\rangle|B'\rangle + \gamma_3|A'\rangle|B\rangle + \gamma_4|A'\rangle|B'\rangle, \quad (66a)$$

where

$$T_1 \otimes T_2 |\Theta\rangle = (\alpha ace^{i(\phi_1+\phi_2)} + \beta bde^{-i(\phi'_1+\phi'_2)})|U_1\rangle|U_2\rangle + (\alpha ade^{i(\phi_1+\phi'_2)} - \beta bce^{-i(\phi'_1+\phi_2)})|U_1\rangle|L_2\rangle + (\alpha bce^{i(\phi'_1+\phi_2)} - \beta ade^{-i(\phi_1+\phi'_2)})|L_1\rangle|U_2\rangle + (\alpha abde^{i(\phi'_1+\phi'_2)} + \beta ace^{-i(\phi_1+\phi_2)})|L_1\rangle|L_2\rangle. \quad (71)$$

We now calculate the probability of joint output into beams U_1 and U_2 (or equivalently of joint detection by ideal detectors placed in these beams), which we shall denote $P(U_1 U_2)$, as well as the analogous probabilities $P(U_1 L_2)$, $P(L_1 U_2)$, and $P(L_1 L_2)$. From these we can calculate the single probabilities $P(U_i)$ and $P(L_i)$ ($i=1,2$),

$$P(U_1 U_2) = \alpha^2 a^2 c^2 + \beta^2 b^2 d^2 + 2\alpha\beta abcd \cos\Phi, \quad (72)$$

where

$$|\gamma_1|^2 + |\gamma_2|^2 + |\gamma_3|^2 + |\gamma_4|^2 = 1. \quad (66b)$$

It is understood that $|\Theta\rangle$ should be symmetrized since photons are bosons, but the results that we obtain without explicit symmetrization would not be changed by writing a symmetrized version of Eq. (66a), provided that the subspace spanned by $|A\rangle, |A'\rangle$ is orthogonal to that spanned by $|B\rangle, |B'\rangle$, and likewise for $|U_1\rangle, |L_1\rangle$ and $|U_2\rangle, |L_2\rangle$.

By the well-known theorem of Schmidt [12], $|\Theta\rangle$ can be expressed as

$$|\Theta\rangle = \alpha|C\rangle|D\rangle + \beta|C'\rangle|D'\rangle, \quad (67a)$$

where $|C\rangle$ and $|C'\rangle$ constitute an orthonormal basis in the subspace spanned by $|A\rangle$ and $|A'\rangle$, while $|D\rangle$ and $|D'\rangle$ constitute an orthonormal basis in the subspace spanned by $|B\rangle$ and $|B'\rangle$. The coefficients α and β can be chosen to be real by using phase options for the vectors $|C\rangle, |C'\rangle, |D\rangle$, and $|D'\rangle$, and

$$\alpha^2 + \beta^2 = 1. \quad (67b)$$

The most general unitary unimodular mapping T_1 relating the specified domain and counterdomain for photon 1 can be expressed in terms of the $|C\rangle, |C'\rangle$ basis as

$$T_1|C\rangle = ae^{i\phi_1}|U_1\rangle + be^{i\phi'_1}|L_1\rangle, \quad (68a)$$

$$T_1|C'\rangle = -be^{-i\phi'_1}|U_1\rangle + ae^{-i\phi_1}|L_1\rangle, \quad (68b)$$

where a and b are real numbers whose squares sum to unity. Likewise,

$$T_2|D\rangle = ce^{i\phi_2}|U_2\rangle + de^{i\phi'_2}|L_2\rangle, \quad (69a)$$

$$T_2|D'\rangle = -de^{-i\phi'_2}|U_2\rangle + ce^{-i\phi_2}|L_2\rangle, \quad (69b)$$

c and d being real numbers whose squares sum to unity. The pair of transducers is represented by

$$T = T_1 \otimes T_2, \quad (70)$$

which is unitary unimodular mapping from the space initially associated with the photon pair $1+2$ into the space of output states. From Eqs. (67)–(70) we obtain

$$\Phi = \phi_1 + \phi'_1 + \phi_2 + \phi'_2, \quad (73)$$

$$P(U_1 U_2) = \alpha^2 a^2 d^2 + \beta^2 b^2 c^2 - 2\alpha\beta abcd \cos\Phi, \quad (74)$$

$$P(L_1 U_2) = \alpha^2 b^2 c^2 + \beta^2 a^2 d^2 - 2\alpha\beta abcd \cos\Phi, \quad (75)$$

$$P(L_1 L_2) = \alpha^2 b^2 d^2 + \beta^2 a^2 c^2 + 2\alpha\beta abcd \cos\Phi, \quad (76)$$

$$P(U_1) = P(U_1 U_2) + P(U_1 L_2) = \beta^2 + a^2(\alpha^2 - \beta^2), \quad (77)$$

$$P(U_2) = P(U_1 U_2) + P(L_1 U_2) = \beta^2 + c^2(\alpha^2 - \beta^2). \quad (78)$$

The single-particle fringe visibilities V_1 and V_2 can be determined by inspection from Eqs. (9), (77), and (78), together with $\alpha \geq \beta$. Clearly,

$$[P(U_1)]_{\max} = [P(U_2)]_{\max} = \alpha^2, \quad (79a)$$

$$[P(U_1)]_{\min} = [P(U_2)]_{\min} = \beta^2. \quad (79b)$$

Hence,

$$V_i = \frac{\alpha^2 - \beta^2}{\alpha^2 + \beta^2} = \alpha^2 - \beta^2 \quad (i=1,2). \quad (80)$$

We note that $P(U_1)$ achieves its maximum and minimum when a has the respective values 1 and 0. When a is unity (hence b is zero), a photon in C will go with certainty into U_1 and a photon in C' will go with certainty into L_1 . When a is zero (hence b is unity), the exit states are reversed. A similar statement can be made concerning photon 2, relating the vectors $|D\rangle$ and $|D'\rangle$ to exit in beams U_2 and L_2 .

We turn now to the two-particle fringe visibility V_{12} . As pointed out in Ref. [10], one cannot capture the intuitive meaning of two-particle fringe visibility by using the attractive definition

$$V_{12} = \frac{[P(U_1 U_2)]_{\max} - [P(U_1 U_2)]_{\min}}{[P(U_1 U_2)]_{\max} + [P(U_1 U_2)]_{\min}} \quad (\text{No}); \quad (81)$$

this expression would yield a nonzero value even if $|\Theta\rangle$ is a product state, for in that case $P(U_1 U_2)$ is the product of $P(U_1)$ and $P(U_2)$, and these vary respectively with T_1 and T_2 . As in Ref. [10] we define a “corrected” joint probability $\bar{P}(U_1 U_2)$ by subtracting the product $P(U_1)P(U_2)$ from $P(U_1 U_2)$ and adding a constant as a compensation against excessive subtraction:

$$\bar{P}(U_1 U_2) = P(U_1 U_2) - P(U_1)P(U_2) + \frac{1}{4}. \quad (82a)$$

By Eqs. (72), (77), and (78),

$$\bar{P}(U_1 U_2) = \alpha^2 \beta^2 \cos \mu \cos \nu + \frac{1}{2} \alpha \beta \sin \mu \sin \nu \cos \Phi + \frac{1}{4}, \quad (82b)$$

where

$$a = \cos \frac{\mu}{2}, \quad b = \sin \frac{\mu}{2}, \quad c = \cos \frac{\nu}{2}, \quad d = \sin \frac{\nu}{2}. \quad (82c)$$

A rationale for the term $\frac{1}{4}$ in Eq. (82a) is the fact that $\frac{1}{4}$ is the least real number s such that $P(U_1 U_2) - P(U_1)P(U_2) + s$ is non-negative for all two-particle vectors of the form of Eq. (66a) and all unitary mappings T_1 and T_2 of the classes under consideration, as can be checked from Eqs. (72), (77), and (78). We now parallel Ref. [10] and define the two-particle fringe visibility V_{12} as

$$V_{12} = \frac{[\bar{P}(U_1 U_2)]_{\max} - [\bar{P}(U_1 U_2)]_{\min}}{[\bar{P}(U_1 U_2)]_{\max} + [\bar{P}(U_1 U_2)]_{\min}}. \quad (83)$$

To find the extrema of $\bar{P}(U_1 U_2)$ we use Eq. (82b) and set partial derivatives to zero: first,

$$0 = \frac{\partial \bar{P}(U_1 U_2)}{\partial \Phi} = -\frac{1}{2} \alpha \beta \sin \mu \sin \nu \sin \Phi. \quad (84)$$

If $\alpha \beta \neq 0$, then Eq. (84) can be satisfied only if one of the two following conditions is satisfied: (i) $\sin \mu \sin \nu = 0$, in which case

$$\frac{1}{4} - \alpha^2 \beta^2 \leq \bar{P}(U_1 U_2) \leq \frac{1}{4} + \alpha^2 \beta^2, \quad (85)$$

or (ii) $\sin \Phi = 0$, in which case

$$\bar{P}(U_1 U_2) = \alpha^2 \beta^2 \cos \mu \cos \nu \pm \frac{1}{2} \alpha \beta \sin \mu \sin \nu + \frac{1}{4}, \quad (86a)$$

and

$$0 = \frac{\partial \bar{P}}{\partial \mu} = -\alpha^2 \beta^2 \sin \mu \cos \nu \pm \frac{1}{2} \alpha \beta \cos \mu \sin \nu, \quad (86b)$$

$$0 = \frac{\partial \bar{P}}{\partial \nu} = -\alpha^2 \beta^2 \cos \mu \sin \nu \pm \frac{1}{2} \alpha \beta \sin \mu \cos \nu. \quad (86c)$$

If $\alpha^2 \beta^2 \neq \frac{1}{2} \alpha \beta$, then Eqs. (86b) and (86c) imply $\cos \mu \sin \nu = \sin \mu \cos \nu = 0$, which is possible only if one of two conditions is satisfied: (iia) $(\mu, \nu) = (m\pi, n\pi)$, with m, n integers, in which case Eq. (85) is again satisfied, or (iib) $(\mu, \nu) = (\pi/2, \pi/2)$, values are all mod π , in which case

$$\bar{P}(U_1 U_2) = \frac{1}{4} \pm \frac{1}{2} \alpha \beta. \quad (87)$$

But

$$\frac{1}{2} \alpha \beta \geq \alpha^2 \beta^2, \quad (88)$$

and therefore a review of all the cases (i), (iia), and (iib) yields

$$[\bar{P}(U_1 U_2)]_{\max} = \frac{1}{4} + \frac{1}{2} \alpha \beta, \quad (89a)$$

$$[\bar{P}(U_1 U_2)]_{\min} = \frac{1}{4} - \frac{1}{2} \alpha \beta. \quad (89b)$$

Note that in the neglected case of $\beta = 0$ these equations continue to hold, as does Eq. (80), because α was assumed $\geq \beta$. It follows that without exception

$$V_{12} = 2\alpha\beta. \quad (90)$$

By Eqs. (80) and (90),

$$\begin{aligned} V_{12}^2 + V_i^2 &= 4\alpha^2 \beta^2 + (\alpha^2 - \beta^2)^2 \quad (i=1,2) \\ &= (\alpha^2 + \beta^2)^2 = 1, \end{aligned} \quad (91)$$

which is the expression for the complementarity of one-particle and two-particle visibilities promised in the Introduction (slightly generalized, since $i = 1$ or 2).

In Ref. [10] a more restricted set of transducers was considered than the class permitted here. Each T_i was taken to consist of a symmetric beam splitter with reflectivity r and transmittivity t both equal to $1/\sqrt{2}$, together with a phase shifter in one beam incident upon the beam splitter. The small letters v_i ($i=1,2$) and v_{12} denote the one-particle and two-particle visibilities under this restriction. It was shown that for a large class [13] of two-particle vectors $|\Theta\rangle$, the inequality

$$v_{12}^2 + v_i^2 \leq 1 \quad (92)$$

holds. Of course, the visibilities V_i and V_{12} are computed by letting the transducers be represented by any unitary unimodular mappings connecting the relevant subspaces, whereas v_i and v_{12} are computed by restricting attention to a subclass of unitary unimodular mappings.

Inequality (92) can be shown to be a corollary of Eq. (91) in the following way. Let x and x' respectively be maximum and minimum values of $\bar{P}(U_1 U_2)$ when one considers the full class of transducers and the subclass of transducers, and let y and y' be the corresponding minimum values. Then $x \geq x'$, $y \leq y'$, and

$$V_{12} - v_{12} = \frac{x-y}{x+y} - \frac{x'-y'}{x'+y'} = \frac{2(xy'-yx')}{(x+y)(x'+y')} \geq 0 . \quad (93)$$

Likewise

$$V_i - v_i \geq 0 . \quad (94)$$

Hence

$$v_{12}^2 + v_i^2 \leq V_{12}^2 + V_i^2 = 1 . \quad (95)$$

There are two advantages in deriving the inequality (92) in this way, as compared with the *ab initio* derivation of the inequality in Ref. [10]. The first is that the inequality is derived for *any* subclass of transducers represented by unitary operators, not just for the special subclass treated in Ref. [10]. The second is that the derivation of Ref. [10] is valid only for a subclass of the two-photon states of Eq. (66a), because of the lacuna noted by Goldstein and discussed in Ref. [13]. In order to fill this lacuna by the method of Ref. [10] some additional lengthy calculations would be needed, which are avoided by the method of the present paper.

IV. DISCUSSION

The chief results of this paper are the derivations of two complementarities. The first relates the distinguishability $D(\mathcal{P})$ between two paths of a particle, when the preparation \mathcal{P} of an ensemble of replicas of the particle is given, to the fringe visibility v_1 when amplitudes from the two paths are combined. This complementarity is formulated both as an inequality (7b) and an equality (7a), where the maximum is taken over the set of preparations \mathcal{P} compatible with a density operator ρ . The second complementarity relates the one-particle fringe visibility to the two-particle fringe visibility in an important class of two-particle preparations. Again there is a formulation of the complementarity as an equality (11b) and an inequality (11a). The visibilities V_{12} and V_1 involved in Eq. (11b) are obtained by calculating the probabilities of joint and single detections as one ranges over the entire class of passive lossless transducers to which the particles are subjected, whereas v_{12} and v_1 are calculated by ranging over subclasses of these transducers.

Several important conceptual issues are bound up with these complementarities. Some of these have been indicated in the body of the paper, but others were discussed briefly or not at all.

It was pointed out in Sec. I that our measure $D(\mathcal{P})$ of

path distinguishability refers to predictions rather than retrodictions. The great advantage of a future-directed concept is the possibility of an experimental check, by appropriate placement of detectors. The importance of distinguishing between predictions and retrodictions in the interpretation of complementarity relations was made at an early date by Heisenberg [14], who wrote the following.

This formulation makes it clear that the uncertainty relation does not refer to the past; if the velocity of the electron is at first known and the position then exactly measured, the position for times previous to the measurement may be calculated. Then for these past times $\Delta p \Delta q$ is smaller than the usual limiting value, but this knowledge of the past is of a purely speculative character, since it can never (because of the unknown momentum caused by the position measurement) be used as an initial condition in any calculation of the future progress of the electron and thus cannot be subjected to experimental verification.

It is also important to emphasize that the quantity $D(\mathcal{P})$ is *distinguishability*, and the suffix “ability,” connoting physical possibility, is crucial. The limitation upon fringe visibility v_1 that is asserted in Eqs. (7a) and (7b) is not imposed by the actual information that the observer has extracted concerning the particles of interest, but in the information that could in principle be extracted within the constraints established by the preparation. The relevance of possible information for interferometry was dramatically demonstrated by an experiment of Zou, Wang, and Mandel [15] in which “the disappearance of the interference pattern here is not the result of a large uncontrollable disturbance...in the spirit of the Heisenberg γ -ray microscope, but simply a consequence of the fact that the two possible photon paths s_1 or s_2 have become distinguishable....The experiment...emphasizes that the state or density operator reflects not only what is known but to an extent also what could be known, in principle, about the photon.” See also the work of Scully and Walther [16].

Although we agree with the foregoing quotation from Zou, Wang, and Mandel, we add that information about the preparation \mathcal{P} of an ensemble is not always exhausted by the density operator ρ that \mathcal{P} determines. As shown in Sec. II, the distinguishability D depends upon \mathcal{P} , and there exist preparations \mathcal{P} and \mathcal{P}' determining the same ρ , but such that $D(\mathcal{P}) \neq D(\mathcal{P}')$. It would be interesting to inquire more generally which properties of an ensemble are determined by the density operator and which are not.

As noted in Eq. (6) the maximum $D(\mathcal{P})$ among all those preparations corresponding to a given ρ —indeed, corresponding to a given $|\rho_{12}|$ —is achieved by $D(|\Theta\rangle)$, where $|\Theta\rangle \in H_1 \otimes H_2$ yields a density operator for particle 1 with the specified $|\rho_{12}|$. This fact shows that an entangled state $|\Theta\rangle$ establishes certain correlations between particles 1 and 2 and these correlations permit the optimum prediction of the path of 1 on the ground of an ob-

servation made upon 2. It is also noteworthy that the detailed characters of H_2 and of $|\Theta\rangle$, over and above the constraint on $|\rho_{12}|$, are irrelevant. All that matters is a certain structure.

Although the two complementarities exhibited in this paper are distinct, they are intimately related. This fact is most clearly seen in

$$|\Psi\rangle = \frac{1}{\sqrt{2}} [|A\rangle|B\rangle + |A'\rangle|B'\rangle], \quad (96)$$

which is a maximally entangled [17] state of 1+2. Clearly $D(|\Psi\rangle)$ is unity since the path of particle 1 can be predicted with certainty after observing whether particle 2 goes into path B or path B' ; hence by the inequality (7b), $v_1=0$. But to someone unfamiliar with entanglement this conclusion is paradoxical, because $|\Psi\rangle$ describes a pure quantum state, with definite phase relations between the term containing $|A\rangle$ and the term containing $|A'\rangle$, and in elementary interferometry a definite phase relation guarantees visible fringes. Of course, this argument is specious, because the definite phase relation holds between two terms not in H_1 but in $H_1 \otimes H_2$ and is responsible for the value unity of the two-particle fringe visibility; by tracing out the variables of particle 2, the phase relation between $|A\rangle$ and $|A'\rangle$ is lost. In the quantum mechanics of entangled states, the state of 1+2 is more than the state of 1 by itself conjoined with the state of 2 by itself and that "more" resides in the phase relation.

Our discussion of the two complementarities was confined to experimental arrangements in which particle 1 propagates only in two beams A and A' and in which there is only one quantum state in each beam. It is obviously desirable to extend our results. Appendix C suggests an extension of the definition of path distinguishability to the case of n beams, each with a single quantum state, but it says nothing about a complementarity between distinguishability and fringe visibility. Furthermore, it is desirable to demonstrate complementarity relations among one-particle, two-particle, . . . , N -particle fringe visibilities when N -particle entangled states are prepared [18]. At present, however, we do not even have a natural definition for N -particle fringe visibility when N is greater than 2.

ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation under Grants Nos. PHY-90-22345 and PHY-93-21992. We are grateful to Professor Sheldon Goldstein, Professor Michael Horne, Professor Leonard Mandel, Professor Yakir Aharonov, Professor Gopathy Murthy, Professor Charles Willis, and Professor Euan Squires for discussions.

APPENDIX A: COMPARISON WITH INFORMATION THEORETICAL UNCERTAINTY RELATIONS

Wootters and Zurek [2], Mittelstaedt, Prieur, and Schieder [19], Lahti, Busch, and Mittelstaedt [20] and others have used information theory to express the "lack

of information" concerning an observable \mathcal{A} when a pure state ϕ is given. In particular, let $\phi \in H_1$ and \mathcal{A} have two orthogonal eigenvectors $|A\rangle$ and $|A'\rangle$, so that the relevant probabilities are

$$\gamma = |\langle \phi | A \rangle|^2 = \text{Prob}(\mathcal{A} = A | \phi), \quad (\text{A1a})$$

$$1 - \gamma = |\langle \phi | A' \rangle|^2 = \text{Prob}(\mathcal{A} = A' | \phi). \quad (\text{A1b})$$

Then Wootters and Zurek [their Eq. (7)] define the lack of information concerning \mathcal{A} in this situation as

$$H(\gamma) = -[\gamma \ln \gamma + (1-\gamma) \ln(1-\gamma)]. \quad (\text{A2})$$

A plausible candidate $d(\gamma)$ for a measure of distinguishability of the values of \mathcal{A} (generalizing "path distinguishability") with a range [0,1] can be defined in terms of $H(\gamma)$:

$$d(\gamma) = 1 - H(\gamma)/\ln 2. \quad (\text{A3})$$

(Mittelstaedt, Prieur, and Schieder [19] introduce this concept with a different notation on their p. 902.) This function has the desirable feature of equaling 1 when γ is 0 or 1, i.e., when the value of \mathcal{A} can be predicted from $|\phi\rangle$ with certainty, and of equaling 0 when $\gamma = 1 - \gamma = \frac{1}{2}$. In fact, this $d(|\phi\rangle)$ is a monotonically increasing function of $D(|\psi\rangle)$, which is our measure of distinguishability in the pure simple case.

There is, however, no obvious way to maintain this monotonic relation when one passes to the mixed simple case. Suppose, as is natural, that we extend d to the mixed simple case by an analog of Eq. (23):

$$d(\{|\psi_i\rangle, w_i\}) = \sum w_i d(|\psi_i\rangle). \quad (\text{A4})$$

It is easy to exhibit two mixtures in which D has the same value and d has different values, and conversely. For instance, for mixture I,

$$w_1 = \frac{1}{2}, \quad |c_1|^2 = |c'_1|^2 = \frac{1}{2},$$

$$w_2 = \frac{1}{2}, \quad |c_2|^2 = 1, \quad |c'_2|^2 = 0.$$

For mixture II,

$$w_1 = 1, \quad |c_1|^2 = \frac{3}{4}, \quad |c'_1|^2 = \frac{1}{4},$$

$$w_2 = 0, \quad |c_2|^2 \text{ arbitrary}, \quad |c'_2|^2 = (1 - |c_2|^2)^{1/2};$$

$$D(\text{mixture I}) = D(\text{mixture II}) = \frac{1}{2},$$

$$d(\text{mixture I}) = \frac{1}{2},$$

$$d(\text{mixture II}) = 0.1990.$$

Consequently, there is no functional relation between D and d in the mixed case if (A4) is used, and we do not know a natural substitute for (A4). Since our D is an intuitively plausible clarification of the concept of path distinguishability, the failure of a functional relation between D and d makes an information theoretical definition of this concept unattractive.

APPENDIX B: EXPERIMENTAL REALIZATION OF UNITARY UNIMODULAR MATRICES

The general form of an SU(2) matrix is

$$S = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}. \quad (\text{B1})$$

S can be written in the convenient form

$$T(r, \phi, \bar{\phi}, \tilde{\phi}) = \begin{pmatrix} e^{i\bar{\phi}} & 0 \\ 0 & e^{i\phi} \end{pmatrix} \begin{pmatrix} t & ir \\ ir & t \end{pmatrix} \begin{pmatrix} e^{i\tilde{\phi}} & 0 \\ 0 & 1 \end{pmatrix} \quad (\text{B2})$$

by letting

$$\begin{aligned} r &= |b|, \quad t = (1 - r^2)^{1/2}, \quad \phi = \arg b - \frac{\pi}{2}, \\ \bar{\phi} &= -\arg a, \quad \tilde{\phi} = -(\phi + \bar{\phi}) = \arg a - \arg b + \frac{\pi}{2}, \end{aligned} \quad (\text{B3})$$

and making use of $\det S = 1$.

The general properties of beam splitters and phase shifters show that the arrangement of Fig. 3 can be made to realize T of Eq. (70) by the insertion of phase shifters ϕ_i and $\bar{\phi}_i$ in the output beams ($i = 1, 2$) and having the phase shifters ϕ_i before the beam splitters yield phase shifts of $\tilde{\phi}_i$. It is interesting to note that one cannot in general achieve $[P(U_1)]_{\max}$ and $[P(U_1)]_{\min}$ of Eqs. (79a) and (79b) with the same beam splitter but different phase shifters.

APPENDIX C: GENERALIZATION OF DISTINGUISHABILITY

We now propose a definition of *distinguishability* for the general case when the particle of interest has n paths to its detector. In this case one cannot always avoid a net loss when betting on the path taken. Thus, in order for $[0,1]$ to remain the range of the distinguishability measure, we proceed as follows. First, as in the case of $n = 2$, we let $D(\mathcal{P})$ be defined by Eq. (1). Now if the procedure of preparation gives no information whatever for preferring one path to another, there is no better strategy than to predict at random, in which case

$$D(\mathcal{P}) = 1/n - (n-1)/n = -(n-2)/n. \quad (\text{C1})$$

On the other hand, there are cases when $D(\mathcal{P})$ has the value one, for example, when there is perfect correlation between the paths of left-going and right-going particles. If we define the new quantity

$$\bar{D}(\mathcal{P}) = \frac{D(\mathcal{P}) + (n-2)/n}{1 + (n-2)/n}, \quad (\text{C2})$$

then clearly $\bar{D}(\mathcal{P})$ has the range $[0,1]$ and in particular has value 1 when $D(\mathcal{P}) = 1$ and 0 when $D(\mathcal{P}) = -(n-2)/n$. $\bar{D}(\mathcal{P})$ is our proposed general measure of path distinguishability. Clearly, in the case when $n = 2$, $\bar{D}(\mathcal{P})$ agrees with our previously proposed measure in Eq. (1).

- [1] N. Bohr, in *Albert Einstein: Philosopher-Scientist*, edited by P. A. Schilpp (The Library of Living Philosophers, Evanston, IL, 1949), p. 216.
- [2] W. K. Wootters and W. H. Zurek, Phys. Rev. D **19**, 473 (1979).
- [3] L. S. Bartell, Phys. Rev. D **21**, 1698 (1980).
- [4] D. M. Greenberger and A. YaSin, Phys. Lett. A **128**, 391 (1988).
- [5] E. Beltrametti and G. Cassinelli, *The Logic of Quantum Mechanics* (Addison-Wesley, Reading, MA, 1981), p. 56.
- [6] B. D'Espagnat, *Conceptual Foundations of Quantum Mechanics* (Benjamin, Menlo Park, CA, 1971), Chap. 6.
- [7] L. Mandel, Opt. Lett. **16**, 1882 (1991).
- [8] D. C. Burnham and D. L. Weinberg, Phys. Rev. Lett. **25**, 84 (1970); C. O. Alley and Y. H. Shih, in *Proceedings of the Second International Symposium on Foundations of Quantum Mechanics in Light of New Technology*, edited by M. Namiki *et al.* (Physical Society of Japan, Tokyo, 1986), p. 47; Y. H. Shih and C. O. Alley, Phys. Rev. Lett. **61**, 2921 (1988); M. Horne and A. Zeilinger, in *Proceedings of the Symposium on the Foundations of Modern Physics*, edited by P. Lahti and P. Mittelstaedt (World Scientific, Singapore, 1985), p. 435; R. Ghosh and L. Mandel, Phys. Rev. Lett. **59**, 1903 (1987); C. K. Hong, Z. Y. Ou, and L. Mandel, *ibid.* **59**, 2044 (1987); J. D. Franson, *ibid.* **62**, 2205 (1989); J. G. Rarity and P. R. Tapster, *ibid.* **64**, 2495 (1990); P. G. Kwiat, W. A. Vercaka, C. K. Hong, H. Nathel, and R. Y. Chiao, Phys. Rev. A **41**, 2910 (1990); M. Horne, A. Shimony, and A. Zeilinger, in *Quantum Coherence*, edited by J. Anandan (World Scientific, Singapore, 1990), p. 356.
- [9] M. A. Horne and A. Zeilinger in *Proceedings of the Symposium on the Foundations of Modern Physics* (Ref. [8]).
- [10] G. Jaeger, M. A. Horne, and A. Shimony, Phys. Rev. A **48**, 1023 (1993).
- [11] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior* (Princeton University Press, Princeton, 1953), pp. 84, 85, 145, and 146.
- [12] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955), pp. 429–434.
- [13] In Ref. [10] the most general quantum state of a pair of photons in beams A and/or A' and beams B and/or B' , respectively, was written in Eq. (21), using vectors $|A\rangle|B\rangle$, $|A\rangle|B'\rangle$, $|A'\rangle|B\rangle$, and $|A'\rangle|B'\rangle$ with complex coefficients. It was argued that the visibilities v_{11}, v_{22}, v_{12} for this state are exactly the same as for a state in which all the coefficients are real, Eq. (21). Sheldon Goldstein (private communication) pointed out a flaw in the algebraic argument for making all phases zero by a basis change. In fact, only three of the four phases can be set to zero by this procedure in the general case. Hence the demonstration of the complementarity relation in the inequalities (30) and (31) of Ref. [10] was not achieved in full generality. The need for the lengthy calculations required to achieve full generality by the method of Ref. [10] is re-

- moved by the present paper.
- [14] W. Heisenberg, *The Physical Principles of Quantum Theory* (Dover, New York, 1930), p. 20.
 - [15] X. Y. Zou, L. J. Wang, and L. Mandel, Phys. Rev. Lett. **67**, 318 (1991).
 - [16] M. O. Scully and H. Walther, Phys. Rev. A **39**, 5229 (1989).
 - [17] A justification for the phrase “maximally entangled” is proposed by A. Shimony, in *Fundamental Problems in Quantum Theory*, edited by D. Greenberger [Ann. N.Y. Acad. Sci. (to be published)].
 - [18] A special case of complementarity in an n -particle system ($n > 2$) is discussed by D. Greenberger, M. Horne, A. Shimony, and A. Zeilinger, Am. J. Phys. **58**, 1131 (1990), p. 1136.
 - [19] P. Mittelstaedt, A. Prieur, and R. Schieder, Found. Phys. **17**, 891 (1987).
 - [20] P. J. Lahti, P. Busch, and P., Mittelstaedt, J. Math. Phys. **32**, 2770 (1991).