

# Quantitative wave-particle duality in multibeam interferometers

Stephan Dürr

JILA, NIST, and University of Colorado, Boulder, Colorado 80309-0440

(Received 13 February 2001; published 18 September 2001)

We propose quantitative measures of wave properties and particle properties for multibeam interferometers. We show that these quantities are connected by a few fundamental inequalities which express wave-particle duality. Our analysis includes which-way detection schemes and quantum erasure. The inequalities derived here are a generalization of similar limits which were studied previously in the special case of two-beam interferometers. In addition, we propose alternative measures that express the available information more efficiently, and we show that these measures fulfill similar inequalities.

DOI: 10.1103/PhysRevA.64.042113

PACS number(s): 03.65.Ta, 03.67.-a

A quantitative version of wave-particle duality was first investigated in 1979 by Wootters and Zurek [1]. They studied interferometers in which one obtains incomplete which-way information, and showed that a surprisingly strong interference pattern can be retained in this case. The easiest way to obtain incomplete which-way information is to use beams of different intensities, e.g., by making one slit larger than the other in a double-slit experiment. The wave properties and particle properties of such a system can be measured by the visibility  $V$  and predictability  $P$ , respectively. Wave-particle duality is expressed by  $P^2 + V^2 \leq 1$ . In recent years, this relation was generalized to which-way detection schemes [2,3] and quantum erasure [4]. A brief historical review on the theoretical and experimental aspects of the subject can be found in Ref. [5]. Surprisingly, all this research dealt exclusively with two-way interferometers. In the outlook of their 1995 paper, Jaeger, Shimony, and Vaidman [2] asked whether similar quantities with similar limits exist in multibeam interferometers, but nobody has given an answer yet.

In this paper, we analyze multibeam interferometers. The key to a quantitative treatment of wave-particle duality lies in finding useful measures of the wave properties and the particle properties. In Sec. I, we present our definitions of such quantities, and show that they obey the same limit as in the two-beam case. Sections II and III extend our analysis to which-way detection schemes and quantum erasure. We prove that all the wave-particle duality inequalities known for two-beam interferometers also hold for our definitions in multibeam interferometers. In Secs. IV and V, we compare the measure of which-way information that proves useful in the present context with the standard measure of information introduced by Shannon [6], and with a measure for quantum information recently introduced by Brukner and Zeilinger [7,8]. Based on this discussion, we propose and analyze alternative measures for the wave and particle properties in Sec. VI.

## I. VISIBILITY AND PREDICTABILITY

### A. Generalized visibility

Consider an  $n$ -beam interferometer, such as the  $n=4$  example shown in Fig. 1. A beam of quantum objects hits a  $n$ -port beam splitter, which splits the incoming beam into  $n$  beams. Next these beams are redirected so that they all im-

pinge on a second  $n$ -port beam splitter that redistributes population between the beams, thus creating  $n$  output beams that exhibit interference. The interfering quantum objects could be almost everything, such as photons, electrons, neutrons, atoms, molecules, etc. For simplicity, we will refer to them as “atoms” in the rest of the paper. The following discussion applies to any  $n$ -port beam splitter independently of how the beam splitter works in detail. We simply treat the beam splitter as a black box that performs a unitary transformation on the  $n$  beams. Note that for any given  $n$ -dimensional unitary transformation, one can construct a corresponding  $n$ -port beam splitter [9].

The first beam splitter in Fig. 1 may have arbitrary splitting ratios, i.e., the intensities of its output beams may differ even if there is only one input beam. We assume, however, that the second beam splitter has a  $1/n$  splitting ratio. Hence the state vector representing an output beam of the second beam splitter has the generic form

$$|b\rangle = \frac{1}{\sqrt{n}} \begin{pmatrix} e^{i\varphi_1} \\ e^{i\varphi_2} \\ \vdots \\ e^{i\varphi_n} \end{pmatrix}, \quad (1.1)$$

where we chose the  $n$  beams in front of the second beam splitter as a basis of states for a matrix representation. The phases  $\varphi_j$  can be varied independently by properly modifying the second beam splitter. Equivalently, and experimen-

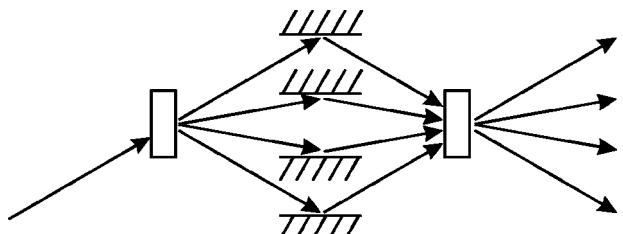


FIG. 1. Scheme of a four-beam interferometer. The incoming beam (left) is split into four beams by a four-port beam splitter. The beams are redirected, e.g., by four individual mirrors, such that all beams impinge on another four-port beam splitter. This beam splitter produces four output beams that exhibit interference.

tally often more easily, this can also be done by inserting  $n$  variable phase shifters—one into each beam in front of the second beam splitter.

The probability that an atom leaves the interferometer in output beam  $|b\rangle$  is

$$I = \langle b | \rho | b \rangle = \frac{1}{n} \sum_{j=1}^n \sum_{k=1}^n \rho_{jk} e^{-i(\varphi_j - \varphi_k)}, \quad (1.2)$$

where  $\rho$  is the statistical operator representing the state of the atoms in front of the second beam splitter. The use of a statistical operator (instead of a state vector) allows us to include possible interactions of the atom with its environment in the formalism.

Using  $\rho^\dagger = \rho$  and  $\text{Tr}\{\rho\} = 1$  (where  $\text{Tr}$  denotes the trace), we can rewrite Eq. (1.2) as

$$I = \frac{1}{n} \left( 1 + \sum_j \sum_{k \neq j} |\rho_{jk}| \cos(\varphi_j - \varphi_k - \arg \rho_{jk}) \right), \quad (1.3)$$

where  $\arg$  denotes the phase of a complex number, i.e.,  $z = |z|e^{i\arg z}$ . Note that  $I$  depends only on the off-diagonal elements of  $\rho$ . In the case of a two-beam interferometer, we obtain

$$I(\varphi_1 - \varphi_2) = \frac{1}{2} [1 + 2|\rho_{12}| \cos(\varphi_1 - \varphi_2 - \arg \rho_{12})]. \quad (1.4)$$

Here the visibility is used as a standard measure for the wave properties of the system. The visibility is usually defined by

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}, \quad (1.5)$$

where  $I_{\max}$  and  $I_{\min}$  denote the maximum and minimum of the interference pattern  $I(\varphi_1 - \varphi_2)$ . This yields

$$V = 2|\rho_{12}|. \quad (1.6)$$

Note that the two-beam interference pattern [Eq. (1.4)] is fully characterized by a single complex number  $\rho_{12}$ . The phase of  $\rho_{12}$  is only referenced to our choice of the coordinate system, and can be set to zero by resetting the zero point of  $\varphi_1 - \varphi_2$ . So all the remaining information about the two-beam interference pattern is contained in  $|\rho_{12}|$ . Precisely this information is expressed by the visibility, as is obvious from Eq. (1.6).

Our aim is to find a quantitative measure of the wave character for  $n$ -beam interferometers. However, definition (1.5) is not necessarily the best starting point for a generalization, because multibeam interference patterns can have many local extrema and their relative position in the  $(\varphi_1, \varphi_2, \dots, \varphi_n)$  parameter space can vary. This variety is due to the large number of off-diagonal matrix elements of  $\rho$  in Eq. (1.3). One-half of  $\rho$ 's  $n(n-1)$  off-diagonal elements is fixed by  $\rho^\dagger = \rho$ , but the other half can vary independently, only limited by the condition  $\rho \geq 0$ . The interference pattern is therefore fully characterized by  $n(n-1)$  real numbers: the modulus and phase of the off-diagonal elements of  $\rho$ . Our choice of the coordinate system allows us to set the zero

points of  $(n-1)$  independent phases,  $\varphi_j$ . (The overall phase of  $|b\rangle$  is irrelevant.) Of course, any measure of the wave character should be independent of our coordinate system. Hence we are left with  $(n-1)^2$  independent real numbers from which to construct a measure of the wave character of the system.

The situation for large  $n$  is somewhat similar to statistical mechanics: We are considering a system with many degrees of freedom and we are not really interested in keeping track of all the details of the system. Instead, we would like to define a “macrovariable” that characterizes global features of the system. This macrovariable should somewhat correspond to our fuzzy concept of the wave character of the system. We therefore propose the following criteria for a measure of the wave properties which we label  $V$ .

(1) It should be possible to give a definition of  $V$  that is based only on the interference pattern  $I$ , without explicitly referring to the matrix elements of  $\rho$ .

(2)  $V$  should vary continuously as a function of the matrix elements of  $\rho$ .

(3) If the system shows no interference (i.e.,  $I = 1/n$  independent of all  $\varphi_j$ ),  $V$  should reach its global minimum.

(4) If  $\rho$  represents a pure state (i.e.,  $\rho^2 = \rho$ ) and all  $n$  beams are equally populated (i.e., all  $\rho_{jj} = 1/n$ ),  $V$  should reach its global maximum.

(5)  $V$  considered as a function in the parameter space  $(\rho_{11}, \rho_{12}, \dots, \rho_{nn})$  should have only global extrema, no local ones.

(6)  $V$  should be independent of our choice of the coordinate system, i.e., insensitive to resetting the zero points of the phases  $\varphi_j$  and insensitive to changing the numbering of the beams.

In the following, we propose and pursue only one definition, although this is clearly not the only possible choice. Instead of focusing on the extrema of the interference pattern, as in Eq. (1.5), rather we consider moments of the function  $I(\varphi_1, \varphi_2, \dots, \varphi_n)$ . We denote the average over all the phases by  $\langle \dots \rangle_\varphi$ , i.e.,

$$\langle f \rangle_\varphi = \frac{1}{(2\pi)^n} \int_0^{2\pi} d\varphi_1 \int_0^{2\pi} d\varphi_2 \cdots \int_0^{2\pi} d\varphi_n f \quad (1.7)$$

for any function  $f(\varphi_1, \varphi_2, \dots, \varphi_n)$ . We easily obtain the first and second moments of the interference pattern [Eq. (1.3)],

$$\langle I \rangle_\varphi = \frac{1}{n}, \quad (1.8)$$

$$\sqrt{\langle (\Delta I)^2 \rangle_\varphi} = \frac{1}{n} \sqrt{\sum_j \sum_{k \neq j} |\rho_{jk}|^2}, \quad (1.9)$$

where  $\Delta I = I - \langle I \rangle_\varphi$  denotes the deviation of  $I$  from its mean value. If the beams are incoherent, we have  $\Delta I = 0$  independent of all phases  $\varphi_j$ . The root mean square (rms) spread  $\sqrt{\langle (\Delta I)^2 \rangle_\varphi}$  therefore is some measure of the amount of interference in the system. Its minimum is zero and its maximum

$\sqrt{(n-1)/n^3}$  is reached if all  $|\rho_{jk}|=1/n$ . We normalize this quantity to define the visibility for  $n$ -beam interferometers:

$$V = \left( \frac{n^3}{n-1} \langle (\Delta I)^2 \rangle_\varphi \right)^{1/2}. \quad (1.10)$$

Thus  $V$  can cover the full range

$$0 \leq V \leq 1. \quad (1.11)$$

Our above definition is based on experimentally accessible features of the interference pattern, just as in Eq. (1.5). We can use Eq. (1.9) to alternatively express the visibility in terms of matrix elements of  $\rho$ , as in Eq. (1.6):

$$V = \left( \frac{n}{n-1} \sum_j \sum_{k \neq j} |\rho_{jk}|^2 \right)^{1/2}. \quad (1.12)$$

In the two-beam case, this equation simply reads  $V = \sqrt{2(|\rho_{12}|^2 + |\rho_{21}|^2)} = 2|\rho_{12}|$ , which is identical to Eq. (1.6). Hence in the two-beam case our definition of the visibility [Eq. (1.10)] is equivalent to the standard definition [Eq. (1.5)].

### B. Generalized predictability

For two-beam interferometers, a useful measure of the particle properties was introduced by Greenberger and YaSin [10]:

$$P = |\rho_{11} - \rho_{22}|. \quad (1.13)$$

They arrived at this definition by considering bets about the atom's way. If we should make a prediction, about which beam an individual atom belongs to, the best strategy would be to bet on the most populated beam. The probability of winning this bet is  $(1+P)/2$ , and the probability of losing it is  $(1-P)/2$ . Hence the average net gain is  $P$ . Englert coined the name "predictability" for  $P$  [3].

In the  $n$ -beam case,  $n-1$  real numbers are required to represent the full information about all the probabilities,  $\rho_{jj}$  that the atom takes one of the ways. Again, we are searching for a "macrovariable" in order to measure the particle character. Our intuitive concept of a particle character is closely related to the concept of information about the ways. We therefore require any measure of the particle properties to fulfill most of the criteria we impose on measures of information. This leads us to the following list, in which we label the measure of the particle properties by  $P$ .

(1)  $P$  should be a continuous function of the probabilities  $\rho_{jj}$ .

(2) If we know the atom's way for sure (i.e.,  $\rho_{jj}=1$  for one beam, implying  $\rho_{jj}=0$  for all other beams),  $P$  should reach its global maximum.

(3) If all ways are equally likely (i.e., all  $\rho_{jj}=1/n$ ),  $P$  should reach its global minimum.

(4) Any change toward equalization of the probabilities  $\rho_{11}, \rho_{22}, \dots, \rho_{nn}$  should decrease  $P$ . Thus if  $\rho_{11} < \rho_{22}$  and we increase  $\rho_{11}$ , decreasing  $\rho_{22}$  an equal amount so that  $\rho_{11}$

and  $\rho_{22}$  are more nearly equal, then  $P$  should decrease. More generally, if we perform any "averaging" operation on the  $\rho_{jj}$  of the form

$$\rho'_{jj} = \sum_k a_{jk} \rho_{kk}, \quad (1.14)$$

where  $\sum_j a_{jk} = \sum_k a_{jk} = 1$ , and all  $a_{jk} \geq 0$ , then  $P$  should decrease (except in the special case where this transformation amounts to no more than a permutation of the  $\rho_{jj}$  with  $P$  of course remaining the same).

(5)  $P$  (together with the resulting quantities defined in Sec. II) should fulfill Eq. (2.5).

Note that criterion (4) implies that  $P$  has only global extrema. Moreover, based on criterion (4), we reject the following measure proposed in Appendix C of Ref. [2]:

$$\frac{n}{n-1} \left( \max_j \{\rho_{jj}\} - \frac{1}{n} \right), \quad (1.15)$$

at which the authors arrived by generalizing the betting strategy discussed by Greenberger and YaSin.

Again, we propose and pursue only one possible definition here. And again, we try the moments of our distribution at hand. We denote the average over the beams by  $\langle \dots \rangle_j$ , i.e.,  $\langle f_j \rangle_j = (1/n) \sum_j f_j$  where  $f_j$  is any set of  $n$  numbers. The mean value of the probabilities  $\rho_{jj}$  is  $\langle \rho_{jj} \rangle_j = 1/n$  because of  $\text{Tr}\{\rho\} = 1$ . The rms spread  $\sqrt{\langle (\rho_{jj} - 1/n)^2 \rangle_j}$  is a possible measure of which-way information that complies with the above criteria. In analogy to the visibility, we define the predictability as the normalized rms spread

$$P = \left[ \frac{n}{n-1} \sum_j \left( \rho_{jj} - \frac{1}{n} \right)^2 \right]^{1/2}. \quad (1.16)$$

We have normalized  $P$  such that it can cover the full range

$$0 \leq P \leq 1. \quad (1.17)$$

Note that using  $\text{Tr}\{\rho\} = 1$ , we can rewrite  $P$  as

$$P = \left[ \frac{n}{n-1} \left( -\frac{1}{n} + \sum_j \rho_{jj}^2 \right) \right]^{1/2}. \quad (1.18)$$

One can easily show that in the two-beam case, our definition of the predictability is equivalent to the standard definition [Eq. (1.13)]. We will discuss alternative measures in Secs. IV and VI.

### C. Wave-particle duality

At a given predictability, the visibility of the interference pattern is limited by

$$P^2 + V^2 \leq 1. \quad (1.19)$$

This inequality is well known for two-beam interferometers (see Ref. [5] and references therein). One of the central results of this paper is that our above defined variables  $V$  and  $P$

obey the same inequality in the  $n$ -beam case. The existence of such a simple relationship between these “macrovariables” qualifies them over many other possible candidates. Equation (1.19) is a quantitative formulation of the concept of wave-particle duality. No matter what the details of our interferometer are, whenever we have a large predictability the visibility of the interference pattern is low. In order to prove Eq. (1.19), we follow an idea from Ref. [11]: we use Eqs. (1.12) and (1.18) and obtain

$$\text{Tr}\{\rho^2\} = \frac{1}{n} + \frac{n-1}{n}(P^2 + V^2). \quad (1.20)$$

The simple fact that  $\text{Tr}\{\rho^2\} \leq 1$  completes the proof of Eq. (1.19). The equality holds in Eq. (1.19) if  $\text{Tr}\{\rho^2\} = 1$ , i.e., if and only if  $\rho$  represents a pure state.

## II. WHICH-WAY DETECTION

So far, we have restricted our considerations to the case where our which-way knowledge is only due to an imbalance in the intensities of the beams. It is, of course, more interesting to investigate schemes that involve which-way detectors. Such a which-way detector is a second quantum system (a part of the atom’s environment) that interacts with the atom. We send only one atom at a time through the interferometer, and always prepare the which-way detector in the same state before the atom enters the interferometer. While the atom is on its way from the first beam splitter to the second beam splitter, we let the atom interact with the which-way detector. The interaction shall be designed such that it changes the state of the which-way detector depending on the way taken by the atom. However, the interaction shall not change the way taken by the atom, thus leaving all the  $\rho_{jj}$  and therefore  $P$  unchanged. Such an interaction stores which-way information in the which-way detector.

For example, let  $|\psi_j\rangle$  denote the state vector if all atoms move through the interferometer along beam  $j$ , and consider the case where prior to the interaction the atoms are in an arbitrary pure state  $|\psi_{\text{in}}\rangle = \sum_j c_j |\psi_j\rangle$  with  $\sum_j |c_j|^2 = 1$ . The interaction could be designed such that it changes the state of the total system (atom plus which-way detector) to

$$\sum_j c_j |\psi_j\rangle \otimes |\chi_j\rangle, \quad (2.1)$$

where the final states of the which-way detector  $|\chi_j\rangle$  depend on  $j$ . Note that  $|\chi_j\rangle$  are normalized but not necessarily orthogonal. In general, the output state is an entangled state. This entanglement (more precisely, the creation of correlations between the which-way detector and the atom’s way) is essential for the storage of which-way information.

To read out the which-way information, we have to perform a measurement of an observable  $W$  on the which-way detector. We denote the eigenvalues of  $W$  by  $w_l$  and the corresponding eigenstates by  $|w_l\rangle$ . We only consider observables with nondegenerate eigenvalues here. We sort the ensemble into subensembles depending on the outcome of the

measurement of  $W$ . The subensemble obtained under the condition that the eigenvalue  $w_l$  was found is described by the statistical operator

$$\rho_{(l)} = \frac{\langle w_l | \tilde{\rho} | w_l \rangle}{p_l}, \quad (2.2)$$

where  $p_l$  is the probability of finding  $w_l$ , and where  $\tilde{\rho}$  denotes the statistical operator of the total system (atom plus which-way detector).

For each subensemble, we define the “conditioned which-way knowledge”  $K_l$  as the predictability of the subensemble in the sense of Eq. (1.16), i.e.,

$$K_l = \left[ \frac{n}{n-1} \sum_j \left( \frac{\langle w_l | \tilde{\rho}_{jj} | w_l \rangle}{p_l} - \frac{1}{n} \right)^2 \right]^{1/2}. \quad (2.3)$$

Note that there is a difference in the time at which the information represented by  $P$  and  $K_l$  is available.  $P$  is only based on the splitting ratio of the first beam splitter, and represents *a priori* which-way knowledge. We have this knowledge before the atom enters the interferometer, and we can use it to literally predict the way.  $K_l$ , however, represents *a posteriori* knowledge. We have to wait until the atom enters the interferometer and interacts with the which-way detector, and then we have to read out the which-way detector, before we obtain this type of knowledge. This knowledge can be used for retrodiction rather than prediction, which is why we use a different name.

For the total ensemble, we define the “which-way knowledge” by averaging over all possible outcomes  $w_l$ :

$$K_W = \sum_l p_l K_l. \quad (2.4)$$

$K_W$  represents the amount of which-way information obtained on average, given that the observable  $W$  was measured. The lower limit of  $K_W$  is

$$P \leq K_W. \quad (2.5)$$

We emphasize that Eq. (2.5) is essential for our interpretation of  $P$  and  $K_W$  in terms of which-way information, because it reflects the fact that our which-way knowledge cannot decrease by reading out the information that is already stored in the which-way detector. Therefore, Eq. (2.5) is on our list of criteria given above.

We now prove that Eq. (2.5) holds with our above definitions. We denote the conditioned probability to find an atom in beam  $j$  conditioned that the which-way detector is found in state  $|w_l\rangle$  by

$$p_{j||l} = \frac{\langle w_l | \tilde{\rho}_{jj} | w_l \rangle}{p_l}. \quad (2.6)$$

We easily obtain

$$K_W^2 = \frac{n}{n-1} \sum_l \sum_m p_l p_m \times \left[ \sum_j \left( p_{j|l} - \frac{1}{n} \right)^2 \right]^{1/2} \left[ \sum_j \left( p_{j|m} - \frac{1}{n} \right)^2 \right]^{1/2}. \quad (2.7)$$

Here we use the Cauchy-Schwarz inequality

$$\left| \sum_j a_j b_j \right| \leq \sqrt{\sum_j a_j^2} \sqrt{\sum_j b_j^2}, \quad (2.8)$$

which holds for all real numbers  $a_j$  and  $b_j$ . We also apply  $x \leq |x|$  (which holds for all real numbers) to obtain  $\sum_j a_j b_j \leq |\sum_j a_j b_j|$ . This yields

$$K_W^2 \geq \frac{n}{n-1} \sum_l \sum_m p_l p_m \sum_j \left( p_{j|l} - \frac{1}{n} \right) \left( p_{j|m} - \frac{1}{n} \right) \quad (2.9)$$

$$= \frac{n}{n-1} \sum_j \left( \text{Tr}_D \{ \tilde{\rho}_{jk} \} - \frac{1}{n} \right)^2 = P^2, \quad (2.10)$$

where  $\text{Tr}_D \{ \tilde{\rho}_{jk} \} = \sum_l \langle w_l | \tilde{\rho}_{jk} | w_l \rangle$  denotes the trace over the which-way detector. This completes the proof of Eq. (2.5).

The which-way knowledge  $K_W$  depends on our choice of the readout observable  $W$ . In order to quantify how much which-way information is actually *stored* in the which-way detector, the arbitrariness of the read-out process can be eliminated by defining the “distinguishability of the ways”

$$D = \max_W \{ K_W \} \quad (2.11)$$

as the maximum of the which-way knowledge  $K_W$  that is obtained at the best choice of the observable  $W$ .

Because of wave-particle duality the storage of which-way information cannot come for free. The correlation created during the interaction with the which-way detector generally reduces the visibility of the interference pattern [12,13]. Trying to quantify this, we are naturally led to one of the most intriguing questions in the context of which-way measurements: How much which-way information can be obtained for a given value of the visibility? The answer is given in form of the “duality relation”

$$D^2 + V^2 \leq 1. \quad (2.12)$$

This inequality generalizes the limit of wave-particle duality set by Eq. (1.19) to the cases including which-way detectors: If we do a very good job on which-way detection, we cannot hope to see a large visibility at the same time. The duality relation is well known in the context of two-beam interferometers [2,3,14–16]. We will prove in Sec. III that it also holds for our definitions in the  $n$ -beam case.

### III. QUANTUM ERASURE

The above discussed loss of visibility due to the storage of which-way knowledge is not irrecoverable. Subensemble sorting as described above offers a possibility to regain the

visibility in carefully chosen subensembles. For reasons discussed below, this strategy is called “quantum erasure” [17].

In analogy to the conditioned which-way knowledge, we define the “conditioned visibility,”  $V_l$ , as the visibility of the subensemble in the sense of Eq. (1.10). This yields

$$V_l = \left( \frac{n}{n-1} \sum_j \sum_{k \neq j} \left| \frac{\langle w_l | \tilde{\rho}_{jk} | w_l \rangle}{p_l} \right|^2 \right)^{1/2}, \quad (3.1)$$

in analogy to Eq. (1.12). We also define the “erasure visibility” as the ensemble average

$$V_W = \sum_l p_l V_l. \quad (3.2)$$

(Note that this quantity was also called “conditioned visibility” [4] or “mean conditioned visibility” [11].) The lower limit of  $V_W$  is

$$V \leq V_W. \quad (3.3)$$

The idea in quantum erasure is to increase the visibility by subensemble sorting. Equation (3.3) guarantees that quantum erasure cannot make things worse on average. The proof of Eq. (3.3) is similar to the proof of Eq. (2.5): we apply the Cauchy-Schwarz inequality to real numbers such as  $|\langle w_l | \tilde{\rho}_{jk} | w_l \rangle / p_l|$ , and obtain

$$V_W^2 \geq \frac{n}{n-1} \sum_j \sum_{k \neq j} \left( \sum_l |\langle w_l | \tilde{\rho}_{jk} | w_l \rangle| \right)^2. \quad (3.4)$$

Now we use a triangular inequality,  $|\sum_j a_j| \leq \sum_j |a_j|$ , and obtain

$$V_W^2 \geq \frac{n}{n-1} \sum_j \sum_{k \neq j} \left| \sum_l \langle w_l | \tilde{\rho}_{jk} | w_l \rangle \right|^2 = V^2. \quad (3.5)$$

In the context of two-way interferometers, Björk and Karlsson [4] made the interesting discovery that the quantities  $K_W$  and  $V_W$  for the same observable  $W$  are limited by an inequality much like Eq. (1.19):

$$K_W^2 + V_W^2 \leq 1. \quad (3.6)$$

This inequality was experimentally tested in Ref. [11]. Note that Björk and Karlsson originally gave a slightly different version of Eq. (3.6) that involved more sophisticated subensemble sorting. However, their version is equivalent to Eq. (3.6), because one can be derived from the other easily, as discussed in Ref. [5]. It is another central result of this paper, that Eq. (3.6) and its corollaries, the duality relation and the erasure relation, also hold for our definitions in the  $n$ -beam case. To show this, we reproduce the proof given in Ref. [11]: We first note that the proof of Eq. (1.19) applies to each subensemble, yielding

$$K_l^2 + V_l^2 \leq 1. \quad (3.7)$$

Equation (3.6) now directly follows from the Cauchy-Schwarz inequality and Eq. (3.7):

$$K_W^2 + V_W^2 = \sum_l \sum_m p_l p_m (K_l K_m + V_l V_m) \quad (3.8)$$

$$\leq \sum_l \sum_m p_l p_m \sqrt{K_l^2 + V_l^2} \sqrt{K_m^2 + V_m^2} \quad (3.9)$$

$$\leq \sum_l \sum_m p_l p_m = 1. \quad (3.10)$$

Equation (3.7) expresses wave-particle duality in quantum erasure: If we choose to measure an observable  $W$  that regains a large visibility for a specific subensemble, we cannot obtain a large amount of which-way information in the same subensemble. Equation (3.6) makes a similar statement for the ensemble average.

Note that the loss of which-way information in quantum erasure is irrecoverable because of the collapse of the wave function associated with the measurement of the observable  $W$ . If we choose an observable that yields a large erasure visibility, then we see from Eq. (3.6) that we inevitably erase most of the which-way knowledge that was stored in the which-way detector. This is why this type of subensemble sorting is called quantum erasure.

The duality relation [Eq. (2.12)] is a simple corollary of Eq. (3.6), because combining the latter with Eq. (3.3) we obtain

$$K_W^2 + V^2 \leq K_W^2 + V_W^2 \leq 1. \quad (3.11)$$

This holds for all observables  $W$ , in particular for the one which maximizes  $K_W$ , which by definition yields  $D$ .

It is also interesting to consider the maximum of  $V_W$ , which is called “coherence.” In some rare cases, it can happen that this maximum does not exist [5], so that it is more accurate to speak of a supremum here:

$$C = \sup_W \{V_W\}. \quad (3.12)$$

This quantity is limited by the “erasure relation” [5]

$$P^2 + C^2 \leq 1, \quad (3.13)$$

which follows from Eqs. (2.5) and (3.6), analogous to the proof of the duality relation.

The erasure relation completes our list of wave-particle duality inequalities that were previously known in two-beam interferometers. With our above definitions, all these inequalities are also valid in multibeam interferometers.

As mentioned earlier, entanglement plays an interesting role in which-way experiments. In our above example [Eq. (2.1)], full which-way information is stored, i.e.,  $D=1$ , if all  $|\chi_j\rangle$  are mutually orthogonal. In this case, the state Eq. (2.1)

is maximally entangled. However,  $D$  is not a measure of entanglement, because a mixed state with statistical operator

$$\sum_j |c_j|^2 |\psi_j\rangle\langle\psi_j| \otimes |\chi_j\rangle\langle\chi_j| \quad (3.14)$$

also yields  $D=1$ , if all  $|\chi_j\rangle$  are mutually orthogonal.  $D$  only measures how much which-way information can be obtained. Subtracting the part that is due to imbalance in the first beam splitter, we might say that  $D^2 - P^2$  is some measure of the correlation between the which-way detector and the atom’s way. But  $D$  and  $P$  are completely insensitive to coherence between the beams. This coherence is measured by  $C$  and  $V$ . Whether a measure of entanglement can be constructed from the above quantities is an interesting question. This might be difficult, because all these quantities are referenced to a preferred basis of the atom, namely the beams inside and behind the interferometer, respectively. Any measure of entanglement (see, e.g., Ref. [18] and references therein), however, should be invariant under all unitary transformations of each subsystem. For example, it has been pointed out that  $D^2 + C^2 - P^2 - V^2 > 0$  is indicative of entanglement [5], but maximizing this sum requires not only a maximally entangled state but also  $P=0$ .

#### IV. SHANNON’S MEASURE OF INFORMATION

We will now compare  $P$  to other measures of information. We start with the standard measure used in information theory which was introduced by Shannon [6]:

$$H = -k \sum_i p_i \ln p_i. \quad (4.1)$$

Here  $k$  is a positive constant, and  $p_i$  denotes the probability for outcome  $i$  of a chance event (corresponding to  $\rho_{ii}$  in our previous notation.) Note that  $H$  measures lack of information, but of course  $-H$  is a measure of information complying with our criteria given in Sec. I B.

Shannon’s measure was discussed excessively in the literature, and it is very useful in a great variety of fields, including statistical mechanics. In two-beam interferometers,  $H$  and the predictability  $P$  are equivalent measures, because they are monotonic functions of each other. But in multibeam interferometers this is not the case, and we owe the reader some justification, why we introduce a new measure  $P$  in this context. (As pointed out in Appendix A of Ref. [2], this problem already arises in two-beam interferometers with a which-way detector, where  $D$  is not a monotonic function of the average conditioned  $H$ .)

The first reason for introducing  $P$  is that we are actually searching for a measure of particle properties. Despite the fact that this measure should have many properties which we also require for measures of information, there is no obvious reason why the measure of particle properties must be an exact measure of information.

The second (more subtle) reason is that there is a fundamental problem in trying to single out one definition of which-way information. In order to understand this, we have

to recall why Shannon preferred  $H$  over all other possible measures of information. He derived  $H$  from the following postulates.

- (1)  $H$  should be continuous in the  $p_i$ .
- (2) If all  $p_i = 1/n$ , then  $H$  should be a monotonic increasing function of  $n$ .
- (3) If a choice is broken down into two successive choices, the original  $H$  should be the weighted sum of the individual values of  $H$ . See Ref. [6] for an example.

The first two postulates leave room for a large variety of definitions. It is the third postulate that is very strong. The problem with this third postulate is that it is applicable only if quantum coherence between the alternatives is absent or irrelevant. This problem arises from the inseparability of quantum phenomena. For example, any attempt to “break down” an interferometer into two successive interferometers (with fewer beams each) will modify the interference pattern.

There is one way to work around the problem of quantum coherence between the alternatives. The freedom of choice of the basis allows us to choose a basis in which  $\rho$  is diagonal. If we calculate  $H$  (with  $p_i = \rho_{ii}$ ) in such a basis, we arrive at the entropy well known from statistical mechanics. However, applying this strategy in order to quantify which-way information is pointless, because the beams in the interferometer already define the preferred basis, relative to which we want to quantify which-way information. For example, if we first diagonalize  $\rho$ , every pure state has zero entropy—completely independent of the probabilities that an atom takes one way or another.

Shannon’s second postulate is also irrelevant in the context of which-way information. Since we cannot “break down” the interferometer into two smaller ones, the scaling with  $n$  is not important.

In addition to his postulates, Shannon listed six interesting properties of  $H$ . For the complete list we refer the reader to Ref. [6]. Here we only mention that Shannon’s properties 1, 2, 4, and 6 are equivalent to the list of criteria we gave in Sec. I. Shannon’s properties 3 and 5 do not have an obvious counterpart in the context of which-way information. This is because usually the which-way detector itself is not an interferometer. Hence there is no such thing as a probability to find the which-way detector in one of its “ways.”

We conclude that Shannon’s measure of information (without diagonalizing  $\rho$ ) is a possible candidate for quantifying which-way information, but Shannon’s second and third postulates are irrelevant in the present context and cannot qualify  $H$  over any other candidate. Furthermore,  $H$  and  $P$  have remarkably similar properties, and only the existence of simple useful expressions, such as  $P^2 + V^2 \leq 1$ , motivates us to prefer  $P$  over  $H$  in the present context.

## V. A MEASURE OF QUANTUM INFORMATION

Brukner and Zeilinger [7,8] recently proposed a measure of quantum information. They also questioned the uniqueness of Shannon’s measure in situations where quantum coherence is essential. For a quantum measurement in which the possible outcomes occur with probabilities  $\vec{p} = (p_1, p_2, \dots, p_n)$ , they proposed

$$I(\vec{p}) = \sum_{i=1}^n \left( p_i - \frac{1}{n} \right)^2 \quad (5.1)$$

as a measure of the information obtained from the measurement. This is obviously identical to  $[(n-1)/n]P^2$  in our notation.

Brukner and Zeilinger went one step further and considered “complete” sets of mutually “unbiased” observables. Two observables are called unbiased (or complementary) if preparing the system in any eigenstate of the one observable, makes all possible outcomes of a measurement of the other observable equally likely. In a  $n$ -dimensional Hilbert space, a set of mutually unbiased observables consists of at most  $n+1$  observables. If such a set of maximum length exists [19,20], it is called complete. The measurement of a complete set of mutually unbiased observables, of course, requires a large number of identically prepared systems, and only one observable is measured on each system. As a measure of the total information obtained in all these measurements, Brukner and Zeilinger proposed

$$I_{\text{total}} = \sum_{j=1}^{n+1} I(\vec{p}_j) = -\frac{1}{n} + \text{Tr}\{\rho^2\}, \quad (5.2)$$

where  $\vec{p}_j$  represents the probabilities of the outcomes of the measurements of the  $j$ th observable. Brukner and Zeilinger pointed out that  $I_{\text{total}}$  is invariant under unitary transformations and still its constituents  $I(\vec{p}_j)$  have an operational significance for the measurement of single observables.

Obviously, wave-particle duality is closely related to this measure of quantum information. Wave-particle duality also deals with unbiased observables, namely, the atom’s way inside and behind the interferometer. Our assumption that the second beam splitter has a  $1/n$  splitting ratio guarantees that these observables are unbiased. Furthermore, a comparison of Eqs. (1.20) and (5.2) yields

$$I_{\text{total}} = \frac{n-1}{n} (P^2 + V^2). \quad (5.3)$$

Clearly,  $[(n-1)/n]P^2$  represents the information obtained from a measurement of the atom’s way inside the interferometer. Equation (5.3) therefore implies that  $[(n-1)/n]V^2$  is equal to the sum of information obtained from measurements of all the remaining  $n$  members of the complete set of mutually unbiased observables. Our definition of  $V^2$  in Eq. (1.10) is based on an integral instead of a sum, but apparently this produces the same result. The following two-beam example will illustrate this.

A two-beam interferometer is described by a  $2 \times 2$  density matrix and therefore equivalent to a spin- $\frac{1}{2}$  system. We choose the coordinate system of the spin  $\frac{1}{2}$  such that a measurement of the  $z$  component of the spin,  $\sigma_z$ , corresponds to a measurement of the atom’s way inside the interferometer. This yields [16]

$$P^2 = \langle \sigma_z \rangle^2. \quad (5.4)$$

Furthermore, we can choose the  $x$  axis of the coordinate system such that  $V^2 = \langle \sigma_x \rangle^2$  [16]. Note that with this choice of the coordinate system  $\langle \sigma_y \rangle = 0$ . We can alternatively rotate the coordinate system around the  $z$  axis by an arbitrary angle, yielding the more general result

$$V^2 = \langle \sigma_x \rangle^2 + \langle \sigma_y \rangle^2. \quad (5.5)$$

In the two-beam case,  $\{\sigma_x, \sigma_y, \sigma_z\}$  is a complete set of mutually unbiased observables. Equations (5.4) and (5.5) illustrate that  $V^2$  corresponds to the information obtained from the measurement of all except one of the observables from the complete set.

Alternatively, we can rotate the observables (instead of the coordinate system) around the  $z$  axis. We thus obtain

$$V^2 = \langle \sigma(\alpha) \rangle^2 + \langle \sigma(\alpha + \pi/2) \rangle^2, \quad (5.6)$$

where  $\sigma(\alpha)$  denotes the spin along an axis that lies in the  $xy$  plane and subtends an angle of  $\alpha$  with the  $x$  axis. Let us now average Eq. (5.6) over  $\alpha$ . The left side is unchanged, because  $V^2$  is independent of  $\alpha$ . On the right side, the two terms yield the same result because of periodicity. Thus we obtain

$$V^2 = 2 \frac{1}{2\pi} \int_0^{2\pi} \langle \sigma(\alpha) \rangle^2 d\alpha. \quad (5.7)$$

This illustrates that, instead of choosing a specific set of mutually unbiased observables and summing over all except one of these observables, as in Eq. (5.5), we can alternatively [as in Eq. (5.7)] average over all observables that are unbiased with respect to  $\sigma_z$ . Both approaches yield the same result. The integral expression for  $V^2$  in Eq. (1.10) is conceptually interesting, because it shows that all possible combinations of the phase shifters  $\varphi_j$  are equally weighted.

## VI. ALTERNATIVE MEASURES FOR WAVE-PARTICLE DUALITY

In the light of the above measure of quantum information, we can rewrite our measures of wave and particle properties. The natural choices are obviously

$$I_P = \frac{n-1}{n} P^2 = -\frac{1}{n} + \sum_j \rho_{jj}^2, \quad (6.1)$$

$$I_V = \frac{n-1}{n} V^2 = \sum_j \sum_{k \neq j} |\rho_{jk}|^2. \quad (6.2)$$

For which-way detection schemes and subensemble sorting, these yield

$$I_{Kl} = \frac{n-1}{n} K_l^2 = -\frac{1}{n} + \sum_j \left( \frac{\langle w_l | \tilde{\rho}_{jj} | w_l \rangle}{p_l} \right)^2, \quad (6.3)$$

$$I_{Vl} = \frac{n-1}{n} V_l^2 = \sum_j \sum_{k \neq j} \left| \frac{\langle w_l | \tilde{\rho}_{jk} | w_l \rangle}{p_l} \right|^2. \quad (6.4)$$

Evidently,  $I_P$  is equivalent to  $P$ , because they are monotonic functions of each other. The same goes for the other three quantities. Thus we did not really add anything to the theory as yet. This changes, however, if we average over the subensembles. We obtain

$$I_{KW} = \sum_l p_l I_{Kl} = \frac{n-1}{n} \sum_l p_l K_l^2, \quad (6.5)$$

$$I_{VW} = \sum_l p_l I_{Vl} = \frac{n-1}{n} \sum_l p_l V_l^2. \quad (6.6)$$

Note that the measures  $I_{KW}$  and  $I_{VW}$  are not equivalent to  $K_W$  and  $V_W$ , even in the two-beam case. We thus lose the advantage of using definitions compatible with all the earlier work in the literature concerning two-beam interferometers. However, the measures  $I_{KW}$  and  $I_{VW}$  express the available information more efficiently, as we will show below.

We first show that our quantities are somewhat related to the old ones:

$$\frac{n-1}{n} K_W^2 \leq I_{KW}, \quad (6.7)$$

$$\frac{n-1}{n} V_W^2 \leq I_{VW}. \quad (6.8)$$

In order to prove Eq. (6.7), we use the Cauchy-Schwarz inequality, which yields

$$\left( \sum_l p_l K_l \right)^2 \leq \left( \sum_l \sqrt{p_l}^2 \right) \sum_l (K_l \sqrt{p_l})^2. \quad (6.9)$$

Using  $\sum_l p_l = 1$ , we obtain Eq. (6.7). In the same way, we can prove Eq. (6.8).

In analogy to  $P \leq K_W$  and  $V \leq V_W$ , our quantities fulfill

$$I_P \leq I_{KW}, \quad (6.10)$$

$$I_V \leq I_{VW}. \quad (6.11)$$

Equation (6.10) shows the compliance of  $I_P$  with criterion (5) from Sec. I B. To prove Eq. (6.10), we again use the Cauchy-Schwarz inequality, from which we obtain

$$\left( \sum_l p_l p_{j|l} \right)^2 \leq \left( \sum_l \sqrt{p_l}^2 \right) \sum_l (p_{j|l} \sqrt{p_l})^2, \quad (6.12)$$

with  $p_{j|l}$  as defined in Eq. (2.6). Summing over  $j$  yields Eq. (6.10). Along the same lines, we can prove Eq. (6.11).

Wave-particle duality is again expressed by a few inequalities; the easiest two of them are

$$I_P + I_V = -\frac{1}{n} + \text{Tr}\{\rho^2\} \leq \frac{n-1}{n}, \quad (6.13)$$

$$I_{Kl} + I_{Vl} = -\frac{1}{n} + \text{Tr}\{\rho_{(l)}^2\} \leq \frac{n-1}{n}, \quad (6.14)$$

with  $\rho_{(l)}$  as defined in Eq. (2.2). These inequalities are, of course, equivalent to  $P^2 + V^2 \leq 1$  and  $K_l^2 + V_l^2 \leq 1$ . However, we can easily obtain a new limit by averaging Eq. (6.14) over  $l$ :

$$I_{KW} + I_{VW} = -\frac{1}{n} + \sum_l p_l \text{Tr}\{\rho_{(l)}^2\} \leq \frac{n-1}{n}. \quad (6.15)$$

We emphasize that this is not equivalent to  $K_W^2 + V_W^2 \leq 1$ . In fact, Eq. (6.15) is more stringent, because we can derive  $K_W^2 + V_W^2 \leq 1$  from Eqs. (6.7), (6.8), and (6.15). Furthermore, we define the maxima (or suprema if necessary)

$$I_D = \max_w \{I_{KW}\}, \quad (6.16)$$

$$I_C = \max_w \{I_{VW}\}, \quad (6.17)$$

and Eqs. (6.10), (6.11), and (6.15) imply

$$I_D + I_V \leq \frac{n-1}{n}, \quad (6.18)$$

$$I_P + I_C \leq \frac{n-1}{n}, \quad (6.19)$$

which are expressions analogous to  $D^2 + V^2 \leq 1$  and  $P^2 + C^2 \leq 1$ .

Thus, with the measures proposed in this section, we arrive at inequalities that are similar to the ones we obtained in the earlier sections. So, what did we gain? First, the choice of definitions in this section is conceptually more natural in light of the measure of quantum information proposed by Brukner and Zeilinger. Second, in Eq. (6.15) we obtained an

equality in addition to the inequality. The advantage of this is easily seen if we consider the special case where  $\tilde{\rho}$  represents a pure state. In this case obviously all  $\text{Tr}\{\rho_{(l)}^2\} = 1$ , and thus

$$I_{KW} + I_{VW} = \frac{n-1}{n} \quad \text{for all pure states.} \quad (6.20)$$

It is important to note that the analogous equality for the traditional measures,  $K_W^2 + V_W^2 = 1$ , does not hold for arbitrary pure states. (See Refs. [4,11] for examples of  $K_W^2 + V_W^2 \neq 1$  for some pure states.) In other words, the quantities  $I_{KW}$  and  $I_{VW}$  express the available quantum information as efficiently as possible, whereas  $K_W$  and  $V_W$  do not.

Equation (6.20) shows that for pure states, the entanglement with the which-way detector does not prevent us from obtaining the maximum possible total information about the atom, no matter which observable  $W$  is measured on the which-way detector. Which-way experiments are a special class of quantum nondemolition (QND) measurements. It might be interesting to analyze other QND schemes in terms of the Brukner-Zeilinger measure of quantum information.

To summarize, we have shown in this paper that the quantities used to characterize wave and particle properties in two-beam interferometers can be generalized to multibeam interferometers. These generalized measures fulfill the same inequalities as in the two-beam case. Furthermore, we proposed alternative measures of the wave and particle properties that fulfill similar inequalities and express the available quantum information more efficiently.

## ACKNOWLEDGMENT

This work was supported by the NSF.

- 
- [1] W.K. Wootters and W.H. Zurek, Phys. Rev. D **19**, 473 (1979).  
[2] G. Jaeger, A. Shimony, and L. Vaidman, Phys. Rev. A **51**, 54 (1995).  
[3] B.-G. Englert, Phys. Rev. Lett. **77**, 2154 (1996).  
[4] G. Björk and A. Karlsson, Phys. Rev. A **58**, 3477 (1998).  
[5] B.-G. Englert and J.A. Bergou, Opt. Commun. **179**, 337 (2000).  
[6] C.E. Shannon, Bell Syst. Tech. J. **27**, 379 (1948).  
[7] C. Brukner and A. Zeilinger, Phys. Rev. Lett. **83**, 3354 (1999).  
[8] C. Brukner and A. Zeilinger, Phys. Rev. A **63**, 022113 (2001).  
[9] M. Reck, A. Zeilinger, H.J. Bernstein, and P. Bertani, Phys. Rev. Lett. **73**, 58 (1994).  
[10] D.M. Greenberger and A. YaSin, Phys. Lett. A **128**, 391 (1988).  
[11] S. Dürr and G. Rempe, Opt. Commun. **179**, 323 (2000).  
[12] M.O. Scully, B.-G. Englert, and H. Walther, Nature (London) **351**, 111 (1991).  
[13] S. Dürr, T. Nonn, and G. Rempe, Nature (London) **395**, 33 (1998).  
[14] S. Dürr, T. Nonn, and G. Rempe, Phys. Rev. Lett. **81**, 5705 (1998).  
[15] P.D.D. Schwindt, P.G. Kwiat, and B.-G. Englert, Phys. Rev. A **60**, 4285 (1999).  
[16] S. Dürr and G. Rempe, Am. J. Phys. **68**, 1021 (2000).  
[17] M.O. Scully and K. Drühl, Phys. Rev. A **25**, 2208 (1982).  
[18] X. Wang and B.C. Sanders, e-print quant-ph/0104011.  
[19] I.D. Ivanovic, J. Phys. A **14**, 3241 (1981).  
[20] W.K. Wootters and B.D. Fields, Ann. Phys. (N.Y.) **191**, 363 (1989).