PHYSICAL REVIEW A VOLUME 39, NUMBER 6 MARCH 15, 1989

Total noise and nonclassical states

Mark Hillery
Department of Physics and Astronomy, Hunter College of the City University of New York, 695 Park Avenue, New York, New York 10021
(Received 23 May 1988)

The total noise of a field state is a measure of the fluctuations of the field amplitude. It is a minimum for coherent states. As the behavior of a state becomes more nonclassical, its total noise increases. This is shown first for several specific types of nonclassical states, among them squeezed and sub-Poissonian states. These results are generalized by using nonclassical distance to measure how nonclassical a field state is. A lower bound for the total noise is derived that is an increasing function of nonclassical distance. From it one can conclude that highly nonclassical states have large amplitude fluctuations.

I. INTRODUCTION

The concept of the total noise of a quantum state was introduced by Schumaker in a discussion of pure states with Gaussian wave functions.[1] She found it to be useful in classifying the states that are produced from the vacuum by systems whose Hamiltonians are sums of quadratic and linear forms in creation and annihilation operators. Such Hamiltonians describe interactions between modes of the electromagnetic field that occur in nonlinear optics. Certain of these Hamiltonians leave the total noise unchanged, while others, in particular those which lead to squeezing,[2] do not.

The total noise of a quantum state of a single mode, whose density matrix is ρ, can be defined in terms of the operators

$$X_1=(a^\dagger+a)/2,\quad X_2=i(a^\dagger-a)/2\ . \tag{1.1}$$

These operators correspond to the real and imaginary parts of the field amplitude, respectively. The total noise, which is a measure of the total fluctuations of the amplitude, is

$$\begin{aligned}T(\rho)&=(\Delta X_1)^2+(\Delta X_2)^2\\&=\langle a^\dagger a\rangle-\langle a^\dagger\rangle\langle a\rangle+\tfrac{1}{2}\ .\end{aligned} \tag{1.2}$$

As was pointed out by Schumaker the total noise is always greater than or equal to $\frac{1}{2}$ and reaches this value only for coherent states.[1] To find the total noise of a multimode state one simply adds the single-mode contributions.

Total noise is a quantity which can be directly measured. One method, suggested by the definition, is to measure $(\Delta X_1)^2$ and $(\Delta X_2)^2$ by means of homodyne detection and to add the results. Another related method is to sweep the phase of the local oscillator in a homodyne detector. The resulting number of counts is proportional to the total noise of the signal. This method is discussed further in Sec. II.

In this paper it will be shown that the total noise of a state is greater the more nonclassical the state is. It will first be demonstrated for some specific kinds of nonclassical states. For example, as the photon statistics of a state become more sub-Poissonian, its total noise increases. Similar results hold as a state becomes more squeezed. A way of making this more systematic is to use the nonclassical distance of a state, δ, to measure how nonclassical it is.[3] As δ increases so does the total noise. It should be noted that the converse is not true. A thermal state, for example, is classical and, therefore, has a nonclassical distance of zero. On the other hand, its total noise can be made arbitrarily large by letting its average photon number increase. The relationship between nonclassical distance and total noise goes only one way; large nonclassical distance implies large total noise but not vice versa.

II. MEASUREMENT OF TOTAL NOISE

Total noise can be determined by performing a phase-averaged homodyne measurement. That is, one combines the signal with a strong local oscillator at the detector and sweeps the phase of the local oscillator. The fluctuations in the number of photocounts are determined by the total noise of the signal. The discussion of this scheme given here will parallel the analysis of homodyne detection done by Mandel.[4]

Suppose that the signal is in a quantum state with density matrix ρ. Upon mixing with the local oscillator, which is in the coherent state $|\alpha\rangle$, the field density matrix becomes $\rho_\alpha=D(\alpha)\rho D(\alpha)^{-1}$, where $D(\alpha)$ is the coherent state displacement operator. If we compute $(\Delta n)^2-\langle\hat{n}\rangle$ for this field we find

$$\begin{aligned}(\Delta n)^2-\langle\hat{n}\rangle=&\,\alpha^2(\langle a^{\dagger 2}\rangle_\rho-\langle a^\dagger\rangle_\rho^2)\\&+\alpha^{*2}(\langle a^2\rangle_\rho-\langle a\rangle_\rho^2)\\&+2|\alpha|^2(\langle a^\dagger a\rangle_\rho-\langle a^\dagger\rangle_\rho\langle a\rangle_\rho)\ ,\end{aligned} \tag{2.1}$$

where terms of order α or lower have been dropped because it is assumed that the local oscillator is strong, and the brackets with the subscript ρ indicate the average with respect to the initial signal state. The average of an operator in the state ρ_α is denoted by angular brackets without a subscript. Setting $\alpha=|\alpha|e^{i\theta}$ and averaging the above expression over θ gives

$$(1/2\pi)\int_0^{2\pi} d\theta[(\Delta n)^2-\langle \hat{n}\rangle]=2|\alpha|^2[T(\rho)-\tfrac{1}{2}] . \quad (2.2)$$

This can now be expressed in terms of the number of photocounts by introducing the overall quantum efficiency of the detector, η. If m is the random variable denoting the number of photocounts we have that

$$\langle m\rangle=\eta\langle \hat{n}\rangle \quad \langle m(m-1)\rangle=\eta^2\langle \hat{n}(\hat{n}-1)\rangle . \quad (2.3)$$

Substituting these results into Eq. (2.2) and noting that $\langle \hat{n}\rangle=|\alpha|^2$ to highest order in α, yields the result

$$T(\rho)-\tfrac{1}{2}=[1/(4\pi\eta\langle m\rangle)]\int_0^{2\pi} d\theta[(\Delta m)^2-\langle m\rangle] . \quad (2.4)$$

This expression allows one to find the total noise from the photocurrent statistics.

III. SPECIFIC EXAMPLES

A nonclassical state is one whose P representation either goes negative or contains derivatives of δ functions. Such states cannot be modeled as classical stochastic fields. Here we would like to consider three varieties of these states: sub-Poissonian, squeezed, and amplitude-squared squeezed. In each case we will see that as the nonclassical attribute of the state increases, so does its total noise. These examples establish the plausibility of the more general result which will be proved in Sec. IV.

Let us first examine sub-Poissonian states. A state is sub-Poissonian if

$$(\Delta n)^2<\langle \hat{n}\rangle , \quad (3.1)$$

where $\hat{n}=a^\dagger a$ and $(\Delta n)^2=\langle(\hat{n}-\langle \hat{n}\rangle)^2\rangle$. Such a state is nonclassical.

In order to relate properties of the number operator to the total noise we begin with

$$\begin{aligned}|\langle(X_1-\langle X_1\rangle)(\hat{n}-\langle \hat{n}\rangle)\rangle|^2 &\le \langle(X_1-\langle X_1\rangle)^2\rangle\langle(\hat{n}-\langle \hat{n}\rangle)^2\rangle \\ &\le (\Delta X_1)^2(\Delta n)^2 , \quad (3.2)\end{aligned}$$

which follows from the Schwarz inequality. The expression on the left-hand side of the above inequality can be expressed as

$$\begin{aligned}&\langle(X_1-\langle X_1\rangle)(\hat{n}-\langle \hat{n}\rangle)\rangle \\ &\quad=(\langle[X_1,\hat{n}]\rangle+\langle\{X_1-\langle X_1\rangle,\hat{n}-\langle \hat{n}\rangle\}\rangle)/2 , \quad (3.3)\end{aligned}$$

where the curly brackets denote the anticommutator. Because both X_1 and $\hat{n}$ are Hermitian the commutator term on the right-hand side of Eq. (3.3) is imaginary and the anticommutator is real. This implies that

$$\begin{aligned}(\Delta X_1)^2(\Delta n)^2 &\ge |\langle(X_1-\langle X_1\rangle)(\hat{n}-\langle \hat{n}\rangle)\rangle|^2 \\ &\ge \tfrac{1}{4}|\langle[X_1,\hat{n}]\rangle|^2 \ge \tfrac{1}{4}\langle X_2\rangle^2 , \quad (3.4)\end{aligned}$$

where we have used $[X_1,\hat{n}]=iX_2$. In a similar fashion one has that

$$(\Delta X_2)^2(\Delta n)^2 \ge (\tfrac{1}{4})\langle X_1\rangle^2 . \quad (3.5)$$

We now move to the total noise. From Eqs. (3.4) and (3.5) we see that

$$\begin{aligned}&(\tfrac{1}{4})(\langle X_1^2\rangle-\langle X_1\rangle^2+\langle X_2^2\rangle-\langle X_2\rangle^2) \\ &\qquad \ge (\tfrac{1}{4})(\langle X_1^2\rangle+\langle X_2^2\rangle)-(\Delta n)^2[(\Delta X_1)^2+(\Delta X_2)^2] . \quad (3.6)\end{aligned}$$

If we note that $\hat{n}=X_1^2+X_2^2-\frac{1}{2}$, then this becomes

$$T=(\Delta X_1)^2+(\Delta X_2)^2 \ge (\langle \hat{n}\rangle+\tfrac{1}{2})/\{4[(\Delta n)^2+\tfrac{1}{4}]\} . \quad (3.7)$$

From this inequality it is clear that for fixed $\langle \hat{n}\rangle$ as $(\Delta n)^2$ decreases, then T must increase. Therefore, as a state becomes more sub-Poissonian ($\Delta n/\langle \hat{n}\rangle$ decreasing) its total noise increases.

Let us now move to squeezing. This example is particularly simple. The uncertainty relation for X_1 and X_2 is[2]

$$\Delta X_1 \Delta X_2 \ge \tfrac{1}{4} , \quad (3.8)$$

which gives for the total noise

$$T \ge (\Delta X_1)^2+1/(4\Delta X_1)^2 . \quad (3.9)$$

The expression on the right-hand side of the above inequality reaches a minimum when $\Delta X_1=\frac{1}{2}$. This value of ΔX_1 marks the edge of the classical region, i.e., if $\Delta X_1<\frac{1}{2}$, then the state is nonclassical. For $0<\Delta X_1\le\frac{1}{2}$, as ΔX_1 decreases T increases. Therefore, the more squeezed the state is, the greater is its total noise. Note that we have considered only squeezing in the X_1 quadrature component, but squeezing in other components leads to identical results.

Finally, consider amplitude-squared squeezing. This is defined in terms of the operators[5,6]

$$Y_1=(a^{\dagger 2}+a^2)/2, \quad Y_2=i(a^{\dagger 2}-a^2)/2 , \quad (3.10)$$

and a state is said to be amplitude-squared squeezed in the Y_1 direction if $(\Delta Y_1)^2<\langle \hat{n}+\frac{1}{2}\rangle$. States which are squeezed in this sense are nonclassical. In order to relate the total noise to the uncertainties in these observables the following inequalities are of use:

$$(\Delta X_1)^2(\Delta Y_1)^2 \ge \tfrac{1}{4}\langle X_2\rangle^2 , \quad (3.11)$$

$$(\Delta X_2)^2(\Delta Y_1)^2 \ge \tfrac{1}{4}\langle X_1\rangle^2 . \quad (3.12)$$

The derivations are similar to that which resulted in Eq. (3.4). If these inequalities are combined with the equation $\langle \hat{n}\rangle+\frac{1}{2}=\langle X_1^2\rangle+\langle X_2^2\rangle$, one finds

$$\begin{aligned}&(\tfrac{1}{4})(\langle \hat{n}\rangle+\tfrac{1}{2})-[(\Delta X_1)^2+(\Delta X_2)^2](\Delta Y_1)^2 \\ &\qquad \le \tfrac{1}{4}[(\Delta X_1)^2+(\Delta X_2)^2] ,\end{aligned}$$

or

$$T \geq (\langle \hat{n} \rangle + \tfrac{1}{2})/[4(\Delta Y_1)^2+1] \; . \quad (3.13)$$

Here again we see that for fixed $\langle \hat{n} \rangle$ as ΔY_1 decreases the total noise must increase. An identical result holds for any other quadrature component of the square of the amplitude.

Surveying these results we see that in each case it was possible to find a lower bound for the total noise which depended on the behavior of some observable of the system. As this behavior became more nonclassical the lower bound for the total noise increased. We now want to generalize this result beyond the specific examples which have been presented in this section.

IV. NONCLASSICAL DISTANCE AND TOTAL NOISE

We shall accomplish this generalization by making use of the concept of nonclassical distance.[3] For a given single-mode density matrix, its nonclassical distance is defined to be

$$\delta = \inf_{\rho_{\text{cl}}} \|\rho - \rho_{\text{cl}}\|_1 \; , \quad (4.1)$$

where the infenum is taken over all classical density matrixes and $\|\cdot\|_1$ denotes the trace norm. The distance δ is a measure of the extent to which the probability distribution for any observable in the state ρ can deviate from the set of probability distributions for that observable in classical states. It is, therefore, an observable-independent measure of how nonclassical a state is. By its definition δ is between 0 and 2, and a δ of order 1 means that a state is highly nonclassical. In Ref. 3 a number of bounds on δ were computed. One which will be of use here is the following: if ρ is a pure state then

$$\delta \leq 2(1 - \sup_{\alpha} \langle \alpha|\rho|\alpha \rangle)^{1/2} \; , \quad (4.2)$$

where $|\alpha\rangle$ is a coherent state. The goal is to find a lower bound for T which is an increasing function of δ. We shall first find such a bound for the case in which ρ is a pure state. It will then be shown that the pure state result also holds if ρ is a mixed state.

Before proceeding let us state the result. Define $x(\delta) \geq 0$ as the solution to the equation

$$[1-(\delta/2)^2]x = 1 - e^{-x} \; . \quad (4.3)$$

A lower bound for the total noise is then given by

$$T(\rho) \geq h(\delta) = (\delta/2)^2 x - (\tfrac{1}{2})[1-(\delta/2)^2]x^2 + \tfrac{1}{2} \; . \quad (4.4)$$

The function $h(\delta)$ is equal to $\frac{1}{2}$ at $\delta=0$ and goes to infinity as δ approaches 2. It is plotted in Fig. 1. Some of its other properties are discussed in Sec. V.

A. Pure state

Assume that ρ is a pure state. We can then express Eq. (4.2) as

$$\sup_{\alpha} \langle \alpha|\rho|\alpha \rangle \leq 1-(\delta/2)^2 = \eta \; . \quad (4.5)$$

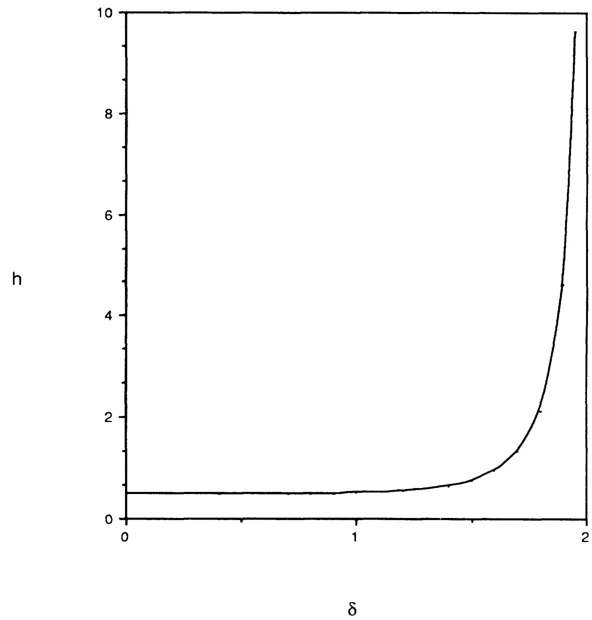

FIG. 1. The function $h(\delta)$ which is a lower bound for the total noise.

This relation tells us that if δ grows, then $s = \sup_{\alpha} \langle \alpha|\rho|\alpha \rangle$ must become smaller. This suggests that we look for a relation between s and the total noise.

This task is made easier if we note that both s and the total noise are invariant under displacement. That is, if $D(\alpha) = \exp(\alpha a^\dagger - \alpha^* a)$, then ρ and $\rho' = D(\alpha)\rho D(\alpha)^{-1}$ have the same value of s and the same total noise. Now choose α so that $\text{Tr}(a\rho')=0$, which implies that $T(\rho') = \text{Tr}(a^\dagger a \rho') + \frac{1}{2}$. Suppose we can prove that for any density matrix ρ_0 which satisfies $\sup_{\alpha}\langle \alpha|\rho_0|\alpha\rangle \leq \eta$ that $\text{Tr}(a^\dagger a \rho_0) \geq \gamma$. Here γ is a constant which will depend on η. Because ρ' satisfies this condition we have that $\text{Tr}(a^\dagger a \rho') \geq \gamma$ which implies that $T(\rho) \geq \gamma + \frac{1}{2}$.

We now want to find γ. To do so we first express $\text{Tr}(a^\dagger a \rho)$ in terms of $\langle \alpha|\rho|\alpha \rangle$,

$$\text{Tr}(a^\dagger a \rho) = (1/\pi) \int d^2\alpha \langle \alpha|\rho|\alpha \rangle |\alpha|^2 - 1 \; . \quad (4.6)$$

Finding γ would seem to come down to minimizing $\int d^2\alpha \langle \alpha|\rho|\alpha \rangle |\alpha|^2$ subject to the constraint $\langle \alpha|\rho|\alpha \rangle \leq \eta$. There are, in fact, two more constraints. The first comes from the normalization condition $\text{Tr}(\rho)=1$ or

$$(1/\pi) \int d^2\alpha \langle \alpha|\rho|\alpha \rangle = 1 \; . \quad (4.7)$$

The second constraint is more complicated. As shown in Appendix A for any $R>0$ we have that

$$\int_{|\alpha| \geq R} d^2\alpha \langle \alpha|\rho|\alpha \rangle \geq \pi e^{-R^2} \; . \quad (4.8)$$

In order to summarize the problem let us define

$$g(r) = (1/\pi) \int_0^{2\pi} d\theta \, r \langle re^{i\theta}|\rho|re^{i\theta} \rangle \; , \quad (4.9)$$

where the coherent state amplitude α has been written as

$\alpha = re^{i\theta}$. Note that $g(r) \geq 0$. Our task is to minimize $B(g) = \int_0^\infty dr\, g(r) r^2$ subject to the constraints

$$g(r) \leq 2r\eta \ , \tag{4.10}$$

$$\int_0^\infty dr\, g(r) = 1 \ , \tag{4.11}$$

$$\int_R^\infty dr\, g(r) \geq e^{-R^2} \ . \tag{4.12}$$

The essential idea behind the solution of this problem is that $g(r)$ is to be chosen as large as possible, in a way consistent with the constraints, for small values of r. This will make $B(g)$ as small as possible because r^2 is a monotonically increasing function. By using this idea as a guide it is possible to give a heuristic argument which leads to the solution of the problem. This will be done here, and the proof that the result is, in fact, correct is left to Appendix B.

First combine Eqs. (4.11) and (4.12) to give

$$\int_0^R dr\, g(r) \leq 1 - e^{-R^2} \ . \tag{4.13}$$

We next need to see whether this constraint or Eq. (4.10) is more restrictive for small values of r. If $g(r)$ satisfies Eq. (4.10), then

$$\int_0^R dr\, g(r) \leq \eta R^2 \ . \tag{4.14}$$

For $\eta < 1$ and sufficiently small R we have that $\eta R^2 < 1 - e^{-R^2}$. Therefore, Eq. (4.10) is the more restrictive bound for small r. This suggests that the $g(r)$ which minimizes $B(g)$ can be constructed in the following fashion. Choose $g_0(r) = 2\eta r$ until Eq. (4.13) is satisfied as an equality. This means that for $0 \leq r \leq r_0$ we have

$$g_0(r) = 2\eta r \ , \tag{4.15}$$

where

$$\eta r_0^2 = 1 - e^{-r_0^2} \ . \tag{4.16}$$

For $r > r_0$ we choose $g_0(r)$ so that Eq. (4.13) is satisfied as an equality. This will make $g_0(r)$ as large as possible for the smaller values of r. That is, for $R > r_0$ we want

$$\int_0^R dr\, g_0(r) = 1 - e^{-r_0^2} + \int_{r_0}^R dr\, g_0(r) = 1 - e^{-R^2} \ . \tag{4.17}$$

Differentiating both sides with respect to R gives

$$g_0(R) = 2Re^{-R^2} \ . \tag{4.18}$$

Summarizing we have

$$g_0(r) = \begin{cases} 2\eta r, & 0 \leq r \leq r_0 \\ 2re^{-r^2}, & r > r_0 . \end{cases} \tag{4.19}$$

Note that this function satisfies Eq. (4.11).

We can now find a lower bound for $B(g)$. For any function satisfying Eqs. (4.10)–(4.12) it must be the case that

$$B(g) \geq B(g_0) = \int_0^{r_0} dr\, 2\eta r^3 + \int_{r_0}^\infty dr\, 2r^3 e^{-r^2} = \eta r_0^4/2 + (1 + r_0^2) e^{-r_0^2} \ . \tag{4.20}$$

This implies that if ρ satisfies $\sup_\alpha \langle \alpha | \rho | \alpha \rangle \leq \eta$, then

$$\mathrm{Tr}(a^\dagger a \rho) \geq B(g_0) - 1 = (1-\eta) r_0^2 - \eta r_0^4/2 \ , \tag{4.21}$$

where Eq. (4.16) has been used to simplify the expression for $B(g_0)$. Finally, the arguments in the paragraph preceding Eq. (4.6) and the above inequality give us

$$T(\rho) \geq f(\eta) = (1-\eta) r_0^2 - \eta r_0^4/2 + \tfrac{1}{2} \ , \tag{4.22}$$

where, it should be noted, r_0 is implicitly determined as a function of η by Eq. (4.16). In terms of the nonclassical distance of the pure state density matrix we have

$$T(\rho) \geq f(1 - (\delta/2)^2) = h(\delta) \ . \tag{4.23}$$

Before discussing this result let us also show that it holds when ρ is a mixed state.

B. Mixed state

In order to demonstrate that our bound is in fact a general one it is necessary to discuss two properties, one of total noise and the other of nonclassical distance. These, and the fact that $h(\delta)$ is a convex function, will allow us to prove our result.

Consider first the total noise. Let ρ_1 and ρ_2 be density matrixes and from them form the density matrix

$$\rho = \theta \rho_1 + (1-\theta) \rho_2 \ , \tag{4.24}$$

where θ is a number between 0 and 1. We want to see how the total noise of ρ is related to that of ρ_1 and ρ_2. Computing $T(\rho)$ we find

$$\begin{aligned} T(\rho) = {} & \theta T(\rho_1) + (1-\theta) T(\rho_2) \\ & + \theta(1-\theta)(\langle a^\dagger \rangle_1 - \langle a^\dagger \rangle_2) \\ & \times (\langle a \rangle_1 - \langle a \rangle_2) \ , \end{aligned} \tag{4.25}$$

where the angular brackets with subscript m indicate expectation values with respect to ρ_m for $m = 1,2$. The last term in Eq. (4.25) is greater than or equal to zero so that

$$T(\rho) \geq \theta T(\rho_1) + (1-\theta) T(\rho_2) \ . \tag{4.26}$$

Finally this result can be generalized to the case in which $\rho = \sum_{n=1}^N \theta_n \rho_n$ where ρ_n is a density matrix and the numbers θ_n satisfy $0 \leq \theta_n \leq 1$ and $\sum_{n=1}^N \theta_n = 1$. We then have

$$T(\rho) \geq \sum_{n=1}^N \theta_n T(\rho_n) \ . \tag{4.27}$$

Now we turn to nonclassical distance. Again consider a density matrix of the form given in Eq. (4.24). If $\rho_{\mathrm{cl}1}$ and $\rho_{\mathrm{cl}2}$ are two classical density matrixes we can use the triangle inequality for norms to show

$$\begin{aligned} & \|\theta \rho_1 + (1-\theta)\rho_2 - \theta \rho_{\mathrm{cl}1} - (1-\theta)\rho_{\mathrm{cl}2}\|_1 \\ & \quad \leq \theta \|\rho_1 - \rho_{\mathrm{cl}1}\|_1 + (1-\theta)\|\rho_2 - \rho_{\mathrm{cl}2}\|_1 \ . \end{aligned} \tag{4.28}$$

Note that if $\rho_{\mathrm{cl}1}$ and $\rho_{\mathrm{cl}2}$ are classical density matrixes, then $\theta \rho_{\mathrm{cl}1} + (1-\theta)\rho_{\mathrm{cl}2}$, where $0 \leq \theta \leq 1$, is also a classical density matrix. This means that

$$\delta=\inf_{\rho_{cl}}\|\rho-\rho_{cl}\|_1\leq\inf_{\rho_{cl1},\rho_{cl2}}\|\rho-\theta\rho_{cl1}-(1-\theta)\rho_{cl2}\|_1\leq\theta\delta_1+(1-\theta)\delta_2\ , \tag{4.29}$$

where δ is the nonclassical distance of ρ and δ_1 and δ_2 are the nonclassical distances of ρ_1 and ρ_2, respectively. This can also be generalized so that if $\rho=\sum_{n=1}^{N}\theta_n\rho_n$, where ρ_n are density matrixes and the numbers θ_n satisfy $0\leq\theta_n\leq 1$ and $\sum_{n=1}^{N}\theta_n=1$, then

$$\delta\leq\sum_{n=1}^{N}\theta_n\delta_n\ . \tag{4.30}$$

In order to make use of these results we need to show that $h(\delta)$ is a convex function, i.e., that for δ_1 and δ_2 greater than or equal to zero and for $0\leq\theta\leq 1$, then

$$h(\theta\delta_1+(1-\theta)\delta_2)\leq\theta h(\delta_1)+(1-\theta)h(\delta_2)\ . \tag{4.31}$$

This follows from the fact that $d^2h/d\delta^2\geq 0$ and is proved in Appendix C. There it is also shown that if $0\leq\theta_n\leq 1$ and $\sum_{n=1}^{N}\theta_n=1$, then

$$h\left[\sum_{n=1}^{N}\theta_n\delta_n\right]\leq\sum_{n=1}^{N}\theta_n h(\delta_n)\ . \tag{4.32}$$

This, in fact, follows from Eq. (4.31). The final fact which we need, that $h(\delta)$ is monotonically increasing, is proved in Appendix C as well.

We can now put all of this together. Let ρ be any density matrix for which $\mathrm{Tr}(a^\dagger a\rho)<\infty$. This density matrix can be diagonalized and expressed in the form

$$\rho=\sum_{n=1}^{\infty}\lambda_n|\psi_n\rangle\langle\psi_n|\ , \tag{4.33}$$

where $\lambda_n\geq 0$ is the nth eigenvalue of ρ and $\sum_{n=1}^{\infty}\lambda_n=1$. Define a sequence of density matrixes ρ_N by

$$\rho_N=(1/\Lambda_N)\sum_{n=1}^{N}\lambda_n|\psi_n\rangle\langle\psi_n|\ , \tag{4.34}$$

with $\Lambda_N=\sum_{n=1}^{N}\lambda_n$. Also let δ_N be the nonclassical distance of ρ_N and δ the nonclassical distance of ρ. As shown in Appendix D $T(\rho_N)$ converges to $T(\rho)$ and δ_N converges to δ. From Eq. (4.27) we have that

$$T(\rho_N)\geq(1/\Lambda_N)\sum_{n=1}^{N}\lambda_n T(|\psi_n\rangle\langle\psi_n|)\ , \tag{4.35}$$

which in conjunction with Eqs. (4.23) and (4.29) implies that

$$T(\rho_N)\geq h\left[(1/\Lambda_N)\sum_{n=1}^{N}\lambda_n\delta_\psi^{(n)}\right]\ , \tag{4.36}$$

where $\delta_\psi^{(n)}$ is the nonclassical distance of $|\psi_n\rangle\langle\psi_n|$. Now making use of the property of nonclassical distance expressed in Eq. (4.30) we find

$$\delta_N\leq(1/\Lambda_N)\sum_{n=1}^{N}\lambda_n\delta_\psi^{(n)}\ . \tag{4.37}$$

Because $h(\delta)$ is a monotonically increasing function we can conclude from Eqs. (4.36) and (4.37) that

$$T(\rho_N)\geq h(\delta_N)\ . \tag{4.38}$$

Letting N go to infinity on both sides gives

$$T(\rho)\geq h(\delta)\ , \tag{4.39}$$

which is the desired result.

V. HIGHLY NONCLASSICAL STATES

The definition of $h(\delta)$ is somewhat complicated but a few features of it can be derived relatively easily. The function achieves its minimum value at $\delta=0$ and is monotonically increasing. If δ is near 2, then η is near zero. An examination of Eq. (4.16) shows that this implies that r_0^2 is large. Therefore, in this regime the $e^{-r_0^2}$ in Eq. (4.16) can be neglected, and $r_0^2\cong 1/\eta$. If this is substituted into the equation for $f(\eta)$ and η is set equal to $1-(\delta/2)^2$, we find

$$h(\delta)\cong(\delta^2/8)/[1-(\delta/2)^2]\ , \tag{5.1}$$

which is valid for δ near 2. As δ approaches 2 $h(\delta)$ and, therefore, the total noise goes to infinity. Highly nonclassical states have very large amplitude fluctuations.

This raises the question of whether there are any states whose nonclassical distance is near 2. First note that for a density matrix ρ the nonclassical distance satisfies

$$\delta=\inf_{\rho_{cl}}\|\rho-\rho_{cl}\|_1\leq\inf_{\rho_{cl}}(\|\rho\|_1+\|\rho_{cl}\|_1)=2\ , \tag{5.2}$$

so that 2 is the largest nonclassical distance possible. Eq. (5.1) implies that any state with $\delta=2$ has an infinite total noise and, therefore, an infinite photon number. Such states are of little physical interest. On the other hand, states whose distance is near 2 will have a large but finite number of photons and are physically reasonable.

In order to show that there are states with nonclassical distance near 2 it is first necessary to develop a new bound for the nonclassical distance of a state. Consider a pure state $\rho=|\psi\rangle\langle\psi|=P_\psi$ for which ρ is identical to the projection onto the state ψ, P_ψ. If ρ_{cl} is a classical density matrix, then

$$\begin{aligned}\|\rho-\rho_{cl}\|_1=\|&P_\psi(\rho-\rho_{cl})P_\psi+P_\perp(\rho-\rho_{cl})P_\psi\\&+P_\psi(\rho-\rho_{cl})P_\perp+P_\perp(\rho-\rho_{cl})P_\perp\|_1\ ,\end{aligned} \tag{5.3}$$

where $P_\perp=I-P_\psi$. Now define the operators A and D to be

$$\begin{aligned}A&=P_\psi(\rho-\rho_{cl})P_\psi+P_\perp(\rho-\rho_{cl})P_\perp\ ,\\D&=P_\perp(\rho-\rho_{cl})P_\psi+P_\psi(\rho-\rho_{cl})P_\perp\ .\end{aligned} \tag{5.4}$$

From the basic properties of norms it follows that

$$|\|A\|_1-\|D\|_1|\leq\|\rho-\rho_{cl}\|_1\leq\|A\|_1+\|D\|_1\ . \tag{5.5}$$

In Ref. 2 it was found that

$$\|A\|_1 = 2(1-\langle\psi|\rho_{cl}|\psi\rangle)\ ,$$

$$\|D\|_1 \le 2(\langle\psi|\rho_{cl}^2|\psi\rangle - \langle\psi|\rho_{cl}|\psi\rangle^2)^{1/2}\ . \tag{5.6}$$

Incorporating Eqs. (5.6) into Eq. (5.5) and taking the infenum over all classical density matrixes gives

$$\delta \le 2\inf_{\rho_{cl}}[(1-\langle\psi|\rho_{cl}|\psi\rangle) + (\langle\psi|\rho_{cl}^2|\psi\rangle - \langle\psi|\rho_{cl}|\psi\rangle^2)^{1/2}]\ ,$$

$$\delta \ge 2\inf_{\rho_{cl}}[(1-\langle\psi|\rho_{cl}|\psi\rangle) - (\langle\psi|\rho_{cl}^2|\psi\rangle - \langle\psi|\rho_{cl}|\psi\rangle^2)^{1/2}]\ , \tag{5.7}$$

where δ is the nonclassical distance of $\rho=|\psi\rangle\langle\psi|$.

Let us now apply this to the photon number state. From Ref. 3 we know that

$$\langle n|\rho_{cl}|n\rangle \le n^n e^{-n}/n! \equiv \gamma_n\ , \tag{5.8}$$

and that $\lim_{n\to\infty}\gamma_n=0$. In order to find a bound for $\langle n|\rho_{cl}^2|n\rangle$, we use the fact that ρ_{cl} has a P representation which behaves like a probability distribution to write

$$\begin{aligned}\langle n|\rho_{cl}^2|n\rangle &= \int d^2\alpha \int d^2\beta \langle n|\alpha\rangle\langle\alpha|\beta\rangle\langle\beta|n\rangle P(\alpha)P(\beta)\\ &= (1/n!)\int d^2\alpha\int d^2\beta e^{-(|\alpha|^2+|\beta|^2-\alpha^*\beta)}\alpha^n\beta^{*n}P(\alpha)P(\beta)\\ &\le \sup_{|\alpha|,|\beta|}(1/n!)|\alpha|^n|\beta|^n e^{-(|\alpha|^2+|\beta|^2)}e^{|\alpha\beta|} \le \gamma_n\ .\end{aligned} \tag{5.9}$$

Substituting these result into the second of Eqs. (5.7) and making use of Eq. (5.2) gives

$$2(1-\gamma_n)-2\sqrt{\gamma_n} \le \delta \le 2\ . \tag{5.10}$$

From this inequality we can see that as n increases the nonclassical distance becomes closer to 2. Therefore, photon number states with large photon numbers have nonclassical distances close to 2.

We can conclude the following. The nonclassical distance of a state lies between zero and two, and there are states with distances throughout this entire range. The inequality given in Eq. (4.39) provides a useful lower bound for the total noise of a state given its nonclassical distance. It implies that the total noise of a highly nonclassical state is large.

VI. CONCLUSION

Total noise is a measure of the size of the amplitude fluctuations of a state of the field. It is a minimum for coherent states. Coherent states are classical. Nonclassical states differ significantly from coherent states in their behavior and the more nonclassical they are the greater the difference. This suggests that the more nonclassical a state is the greater its total noise will be.

This was examined for several examples and found to be true. The actual results took the form of lower bounds on the total noise which are increasing functions of the nonclassical behavior (squeezing, for example). It was then possible to generalize these results by using nonclassical distance, δ, to measure how nonclassical a state is. A lower bound for the total noise was found which is an increasing function of nonclassical distance. The bound goes to infinity as δ approaches its maximum value of 2. This implies that states which are highly nonclassical have very large amplitude fluctuations.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation under Grant No. PHY-8613463 and by a grant from the PSC-CUNY Research Award Program.

APPENDIX A

In this appendix we want to show that for any density matrix ρ

$$\int_{|\alpha|\ge R} d^2\alpha\langle\alpha|\rho|\alpha\rangle \ge \pi e^{-R^2}\ . \tag{A1}$$

We begin by expanding ρ in terms of number states so that

$$\begin{aligned}\int_{|\alpha|\ge R} d^2\alpha\langle\alpha|\rho|\alpha\rangle &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\int_{|\alpha|\ge R} d^2\alpha\langle\alpha|n\rangle\rho_{nm}\langle m|\alpha\rangle\\ &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\int_{|\alpha|\ge R} d^2\alpha\, e^{-|\alpha|^2}(\alpha^{*n}\alpha^m/\sqrt{n!m!})\rho_{nm}\\ &= 2\pi\sum_{n=0}^{\infty}\rho_{nn}\int_R^\infty dr\, e^{-r^2}r^{2n+1}/n!\ ,\end{aligned} \tag{A2}$$

where $r=|\alpha|$. Equation (A2) and the relation $\sum_{n=0}^{\infty}=\rho_{nn}=1$ allows us to derive the lower bound

$$\int_{|\alpha|\geq R} d^2\alpha\langle\alpha|\rho|\alpha\rangle \geq 2\pi \inf_n \left[\int_R^{\infty} dr\, e^{-r^2} r^{2n+1}/n!\right] . \quad \text{(A3)}$$

The integrals on the right-hand side of Eq. (A3) can be evaluated and give

$$\int_R^{\infty} dr\, e^{-r^2} r^{2n+1}/n! = (\tfrac{1}{2})e^{-R^2} \sum_{l=0}^{n} R^{2l}/l! \quad \text{(A4)}$$

from which we can conclude

$$\inf_n \int_R^{\infty} dr\, e^{-r^2} r^{2n+1}/n! = (\tfrac{1}{2})e^{-R^2} . \quad \text{(A5)}$$

Substitution of this result into Eq. (A3) yields Eq. (A1).

APPENDIX B

We now want to prove that the function which minimizes $B(g)$ subject to the constraints given in Eqs. (4.10)–(4.12) is, in fact, $g_0(r)$ as given in Eq. (4.19). In order to do this it is useful to first prove the following lemma.

Lemma: Let $f_1(r)$ and $f_2(r)$ be two piecewise continuous functions which satisfy $f_1(r)\geq 0$, $f_2(r)\geq 0$, and

$$\int_0^{\infty} dr\, f_1(r) = \int_0^{\infty} dr\, f_2(r) = 1 . \quad \text{(B1)}$$

Furthermore, suppose that for all $R>0$

$$\int_0^R dr\, f_1(r) \geq \int_0^R dr\, f_2(r) . \quad \text{(B2)}$$

It then follows that

$$\int_0^{\infty} dr\, r^2 f_1(r) \leq \int_0^{\infty} dr\, r^2 f_2(r) . \quad \text{(B3)}$$

In order to prove this first define the functions

$$F_j(R) = \int_0^R dr\, f_j(r) , \quad \text{(B4)}$$

for $j=1,2$. Equation (B2) implies that

$$\int_0^R dr\, F_1(r) d(r^2)/dr \geq \int_0^R dr\, F_2(r) d(r^2)/dr , \quad \text{(B5)}$$

which upon integration by parts becomes

$$R^2 F_1(R) - \int_0^R dr\, r^2 f_1(r) \geq R^2 F_2(R) - \int_0^R dr\, r^2 f_2(r) . \quad \text{(B6)}$$

If it can be shown that $\lim_{R\to\infty} R^2[F_1(R)-F_2(R)]=0$, then the lemma will be proved.

That this limit vanishes can be demonstrated by first noting that

$$\lim_{R\to\infty} R^2[1-F_2(R)] \geq \lim_{R\to\infty} R^2[F_1(R)-F_2(R)] \geq 0 . \quad \text{(B7)}$$

We can then use the fact that $\int_0^{\infty} dr\, r^2 f_2(r) < \infty$ to conclude

$$\lim_{R\to\infty} \int_R^{\infty} dr\, r^2 d[F_2(r)-1]/dr = 0 . \quad \text{(B8)}$$

This integral provides a bound for the quantity on the left-hand side of Eq. (B7) for if one keeps in mind that $d[F_2(r)-1]/dr \geq 0$, one finds

$$\int_R^{\infty} dr\, r^2 d[F_2(r)-1]/dr \geq R^2 \int_R^{\infty} dr\, d[F_2(r)-1]/dr = R^2[1-F_2(R)] . \quad \text{(B9)}$$

If the $R\to\infty$ limit of both sides is taken and note is taken of Eq. (B8) and the fact that $R^2[1-F_2(R)]\geq 0$, then we see that

$$\lim_{R\to\infty} R^2[1-F_2(R)] = 0 . \quad \text{(B10)}$$

Equation (B7) now allows us to conclude that $\lim_{R\to\infty} R^2[F_1(R)-F_2(R)]=0$. Finally, taking the $R\to\infty$ limit in Eq. (B6) yields Eq. (B3) and proves the lemma.

The application of the lemma to the problem at hand is direct. Suppose that $g(r)$ is a function satisfying Eqs. (4.10)–(4.12). For $0\leq r\leq r_0$ we clearly have that $g_0(r)\geq g(r)$ which gives us that for $0\leq R\leq r_0$

$$\int_0^R dr\, g_0(r) \geq \int_0^R dr\, g(r) . \quad \text{(B11)}$$

If $R>r_0$, then by its definition

$$\int_0^R dr\, g_0(r) = 1-e^{-R^2} \geq \int_0^R dr\, g(r) . \quad \text{(B12)}$$

Therefore, $g_0(r)$ and $g(r)$ correspond to $f_1(r)$ and $f_2(r)$, respectively, in the lemma, and we have that $B(g)\geq B(g_0)$. This means that function $g_0(r)$ gives the minimum value of $B(g)$ subject to the stated conditions.

APPENDIX C

The functions $f(\eta)$ and $h(\delta)$ are defined by Eqs. (4.22) and (4.23), respectively. Several properties of these functions were used to demonstrate that the total noise of a state increases with its nonclassical distance. In particular the fact that $h(\delta)$ is a monotonically increasing, convex function was essential. We now show that this is indeed the case.

Let us first summarize the definitions of $f(\eta)$ and $h(\delta)$,

$$f(\eta) = (1-\eta)x - \eta x^2/2 + \tfrac{1}{2} , \quad \text{(C1)}$$

$$h(\delta) = f(\eta(\delta)) = f(1-(\delta/2)^2) , \quad \text{(C2)}$$

where we have set $x=r_0^2$ and x is determined as a function of η by the condition

$$\eta x = 1-e^{-x} . \quad \text{(C3)}$$

From these equations we want to find $dh/d\delta$ and $d^2h/d\delta^2$.

To begin we differentiate both sides of Eq. (C3) with respect to η to give

$$dx/d\eta = x/(e^{-x}-\eta) = x^2/[e^{-x}(x+1)-1] \ . \quad \text{(C4)}$$

From this equation and the derivative of Eq. (C1) one has that

$$df/d\eta = -x^2/2 \ . \quad \text{(C5)}$$

Application of the chain rule gives

$$dh/d\delta = \delta x^2/4 \ . \quad \text{(C6)}$$

The right-hand side is clearly positive which implies that $h(\delta)$ is an increasing function of δ.

Continuing, differentiate both sides of Eq. (C6) with respect to δ

$$\begin{aligned} d^2h/d\delta^2 &= \tfrac{1}{4}[x^2+2\delta x(dx/d\delta)] \\ &= (x^2/4)\{1-\delta^2 x/[e^{-x}(x+1)-1]\} \ . \end{aligned} \quad \text{(C7)}$$

Noting that $x \ge 0$ implies that $e^{-x}(x+1) \le 1$, we see that $d^2h/d\delta^2 \ge 0$.

It remains to show that these properties of the derivatives of h imply that it is convex. In order to do so assume that $\delta_1 \le \delta_2$ and let θ be a number between 0 and 1. Consider the quantities

$$\begin{aligned} H_1 &= \theta h(\delta_1)+(1-\theta)h(\delta_2) \\ &= h(\delta_1)+(1-\theta)\int_{\delta_1}^{\delta_2} d\delta(dh/d\delta) \ , \end{aligned} \quad \text{(C8)}$$

$$\begin{aligned} H_2 &= h(\theta\delta_1+(1-\theta)\delta_2) \\ &= h(\delta_1)+\int_{\delta_1}^{(1-\theta)(\delta_2-\delta_1)+\delta_1} d\delta(dh/d\delta) \ . \end{aligned} \quad \text{(C9)}$$

Define the function $v(\delta)=(dh/d\delta)|_{\delta_1+\delta}$. Now subtracting Eq. (C9) from Eq. (C8) we find

$$H_1-H_2 = (1-\theta)\int_0^{\delta_2-\delta_1} d\delta\, v(\delta) - \int_0^{(1-\theta)(\delta_2-\delta_1)} d\delta\, v(\delta) \ . \quad \text{(C10)}$$

Changing the variable of integration in the first integral to $(1-\theta)\delta$ gives

$$H_1-H_2 = \int_0^{(1-\theta)(\delta_2-\delta_1)} d\delta[v(\delta/(1-\theta))-v(\delta)] \ . \quad \text{(C11)}$$

The fact that $d^2h/d\delta^2 \ge 0$ implies that $dv/d\delta \ge 0$. Therefore, because $\delta/(1-\theta) > \delta$, the integral in Eq. (C11) is greater than or equal to zero. This, in turn, means that $H_1 \ge H_2$ or

$$h(\theta\delta_1+(1-\theta)\delta_2) \le \theta h(\delta_1)+(1-\theta)h(\delta_2) \ . \quad \text{(C12)}$$

Finally, let us demonstrate the result stated in Eq. (4.32). This can be proved by induction. Equation (C12) shows that it is true for the case $N=2$. Suppose now that it is true for N. We want to show it is true for $N+1$. Let $\{\theta_n | n=1,2,\ldots,N+1\}$ be a set of numbers each of which is between 0 and 1, and which satisfy $\sum_{n=1}^{N+1}\theta_n = 1$. We then have, setting $\Gamma = \sum_{n=1}^{N}\theta_n$,

$$\begin{aligned} h\left[\sum_{n=1}^{N+1}\theta_n\delta_n\right] &= h\left[\Gamma\sum_{n=1}^{N}(\theta_n/\Gamma)\delta_n+\theta_{N+1}\delta_{N+1}\right] \\ &\le \Gamma h\left[\sum_{n=1}^{N}(\theta_n/\Gamma)\delta_n\right]+\theta_{N+1}h(\delta_{N+1}) \ . \end{aligned} \quad \text{(C13)}$$

The above inequality follows by application of Eq. (C12) upon setting $\theta=\Gamma$, and identifying δ_1 in Eq. (C12) with $\sum_{n=1}^{N}(\theta_n/\Gamma)\delta_n$. Because we have assumed that Eq. (4.32) holds for N, it follows from Eq. (C13) that

$$\begin{aligned} h\left[\sum_{n=1}^{N+1}\theta_n\delta_n\right] &\le \Gamma\sum_{n=1}^{N}(\theta_n/\Gamma)h(\delta_n)+\theta_{N+1}h(\delta_{N+1}) \\ &\le \sum_{n=1}^{N+1}\theta_n h(\delta_n) \ , \end{aligned} \quad \text{(C14)}$$

which is the desired result.

APPENDIX D

The density matrixes ρ_N were defined by Eq. (4.34). We want to show that $T(\rho_N)\to T(\rho)$ and $\delta_N \to \delta$ as N goes to infinity.

Let us first examine the total noise. By assumption $\mathrm{Tr}(\rho\hat{n})<\infty$ which implies that $|\mathrm{Tr}(\rho a)|<\infty$. We also have that $\lim_{N\to\infty}\Lambda_N = 1$. This means that for any ϵ we can find an M such that

$$\begin{aligned} &\sum_{n=M+1}^{\infty}\langle\psi_n|\hat{n}|\psi_n\rangle\lambda_n < \epsilon, \quad (1/\Lambda_M)-1<\epsilon \\ &\sum_{n=M+1}^{\infty}|\langle\psi_n|a|\psi_n\rangle|\lambda_n < \epsilon. \end{aligned} \quad \text{(D1)}$$

Looking first at the expectation values of the number operator we have that

$$|\mathrm{Tr}(\rho_M\hat{n})-\mathrm{Tr}(\rho\hat{n})| \le \left|\sum_{n=1}^{M}[(1/\Lambda_M)-1]\langle\psi_n|\hat{n}|\psi_n\rangle\lambda_n - \sum_{n=M+1}^{\infty}\langle\psi_n|\hat{n}|\psi_n\rangle\lambda_n\right| \le \epsilon[\mathrm{Tr}(\rho\hat{n})+1] \ . \quad \text{(D2)}$$

For the expectation values of the annihilation operator

$$|\mathrm{Tr}(\rho_M a)-\mathrm{Tr}(\rho a)| \le \left|\sum_{n=1}^{M}[(1/\Lambda_M)-1]\langle\psi_n|a|\psi_n\rangle\lambda_n - \sum_{n=M+1}^{\infty}\langle\psi_n|a|\psi_n\rangle\lambda_n\right| \le \epsilon\left[\sum_{n=1}^{\infty}|\langle\psi_n|a|\psi_n\rangle|\lambda_n+1\right] \ . \quad \text{(D3)}$$

The right-hand sides of Eqs. (D2) and (D3) can be made arbitrarily small so that $T(\rho_N)\rightarrow T(p)$ as $N\rightarrow\infty$.

Now we consider the nonclassical distances. By arguments similar to those above one can show that

$$\lim_{N\rightarrow\infty} \|\rho_N-\rho\|_1=0\ , \tag{D4}$$

so that we can find an M such that if $N\geq M$, then $\|\rho_N-\rho\|_1<\epsilon$. It then follows that for $N\geq M$ and $\rho_{\rm cl}$ a classical density matrix that

$$\begin{aligned}\|\rho-\rho_{\rm cl}\|_1 &\leq \|\rho-\rho_N\|_1+\|\rho_N-\rho_{\rm cl}\|_1\\ &\leq \|\rho_N-\rho_{\rm cl}\|_1+\epsilon\ ,\\ \|\rho_N-\rho_{\rm cl}\|_1 &\leq \|\rho_N-\rho\|_1+\|\rho-\rho_{\rm cl}\|_1\\ &\leq \|\rho-\rho_{\rm cl}\|_1+\epsilon\ .\end{aligned} \tag{D5}$$

Taking the infenum of both sides of these inequalities over all classical density matrixes gives

$$|\delta-\delta_N|<\epsilon\ , \tag{D6}$$

so that $\lim_{N\rightarrow\infty}\delta_N=\delta$.

[1]B. L. Schumaker, Phys. Rep. **135**, 318 (1986).
[2]For a review see D. F. Walls, Nature **306**, 141 (1983).
[3]M. Hillery, Phys. Rev. A **35**, 725 (1987).
[4]L. Mandel, Phys. Rev. Lett. **49**, 136 (1982).
[5]M. Hillery, Opt. Commun. **62**, 135 (1987).
[6]M. Hillery, Phys. Rev. A **36**, 3796 (1987).