Preparation, measurement and information capacity of optical quantum states

Y. Yamamoto

Musashino Electrical Communication Laboratories, Nippon Telegraph and Telephone Corporation, Musashinoshi, Tokyo 180, Japan

H. A. Haus

Department of Electrical Engineering and Computer Science and Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

The preparation, or generation of coherent states, squeezed states, and photon number states is discussed. The quantum noise is evaluated for various simultaneous measurements of two quadrature components: heterodyning, the beam splitter followed by two single quadrature measurements, the parametric amplifier, the (degenerate and/or nondegenerate) four-wave mixer, the Brillouin and Raman amplifiers, and the laser amplifier. A quantum nondemolition measurement followed by a measurement of the conjugate variable is also categorized as a simultaneous measurement. It is shown that, for all of these schemes, the minimum uncertainty product of the measured variables is exactly equal to that required for a simultaneous measurement of two noncommuting variables. On the other hand, measurements of a single quadrature component are noise-free. Such measurements are degenerate heterodyning, degenerate parametric amplification, and cavity degenerate four-wave mixing and photon counting by a photomultiplier or avalanche photodiode. The Heisenberg uncertainty principle and the quantum-mechanical channel capacity of Shannon are discussed to address the question "How much information can be transmitted by a single photon?" The quantum-mechanical channel capacity provides an upper bound on the achievable information capacity and is ideally realized by photon number states and photon counting detection. Its value is $\hbar\omega/(\ln 2)kT$ bit per photon. The use of coherent or squeezed states and a simultaneous measurement of two quadrature field components or the measurement of one single quadrature field component does not achieve the ultimate limit.

CONTENTS

I. INTRODUCTION

The role of the Heisenberg uncertainty principle as the quantum-mechanical limit on the precision of measurements has been discussed for many years, because it is a crucial issue for the interpretation of quantum mechanics (Bohr, 1935; Einstein *et al.*, 1935; Bohm, Chaps. 5 and 22, 1951; Aharanov and Bohm, 1961, 1964).

Recently, the interest has been rekindled in the two fields, in which measurement accuracies close to the quantum limit imposed by the Heisenberg uncertainty principle have been achieved. The development of the superconducting quantum interferometric device (SQUID) amplifier, which is close to quantum limited sensitivity, raises the possibility of detecting gravitational waves using a mechanically resonant (Weber bar) antenna (Thorne, 1980a, 1980b). Unfortunately, a resonant bar antenna couples so weakly to gravitational waves that the standard quantum limit, imposed by the position-momentum uncertainty relation, $\Delta x \Delta p \geq h/2$, of the bar itself, prevents detection of the excited oscillation of the order of 10^{-19} cm. If the position is measured with an accuracy of $\Delta x \simeq 10^{-19}$ cm, the back action on the momentum $\Delta p \simeq \hbar/2\Delta x$ changes the position by $\Delta x \simeq \Delta v t \simeq \Delta p t/m$ at the instant of the second measurement after t sec. The value is 5×10^{-19} cm when the second measurement is done at $t \simeq 10^{-3}$ sec, even if the mass of the antenna is 10^4 kg.

The new technique called "quantum nondemolition (QND) measurement" was proposed by Braginsky *et al.* (1977) and Unruh (1978) to overcome the standard quantum limit. A similar technique called "back action evading" was proposed by Thorne *et al.* (1978). In the two measurement schemes of the occupation number or of one quadrature component of the mechanical harmonic oscillator the back action of the first measurement is imposed on the conjugate observable of the measured variable. When the subsequent measurement is performed, the back action does not disturb the free motion of the measured variable. This is indeed possible by a proper choice of the measured observable and the interaction (Braginsky *et al.*, 1980; Caves *et al.*, 1980). The other new technique called "contractive state measurement" was proposed by Yuen (1983). In this measurement scheme of the position of a free mass, the back action imposed on the momentum has a quantum-mechanical correlation with the position operator and contributes to the position uncertainty reduction at the instant of the second measurement. A practical scheme to realize this was proposed by Bondurant and Shapiro (1984).

Another field of interest is the measurement of optical fields and the transmission of information using a coherent laser radiation field. The precise measurement of inertial rotation has now become possible by the advent of the four-frequency ring-laser gyroscope (Dorschner *et al.*, 1980) and the ring-fiber gyroscope (Sanders *et al.*, 1980). The sensitivities of these devices have already reached the quantum limit determined by the photon number-phase uncertainty relation, $\Delta n \Delta\varphi \geq \frac{1}{2}$, of the radiation field. The same quantum limit is now being approached by optical communication using coherent heterodyning detection (Yamamoto, 1980; Chan, 1981) and by optical communication using photon counting (Pierce *et al.*, 1981). The quantum limit of the photon number and phase measurements are $\Delta n \simeq \sqrt{\langle n \rangle}$ and $\Delta\varphi \simeq 1/(2\sqrt{\langle n \rangle})$. This standard quantum limit is attributable to the quantum noise of a coherent state (Glauber, 1963).

Quantum states of electromagnetic waves called "squeezed states (Takahashi, 1965; Stoler, 1971; Yuen, 1976)," "photon number states," and "amplitude-squeezed states (or number-phase minimum uncertainty states) (Carruthers and Neito, 1968)" were proposed to bypass the standard quantum limit. These states of the electromagnetic field have reduced quantum noise for one observable and preserve the Heisenberg uncertainty relation by an increased quantum noise for the conjugate observable. The information can be extracted from the observable with reduced quantum noise, and thus the standard quantum limit can be overcome.

In the same way as the Heisenberg uncertainty principle sets an upper limit on the precision of a quantum measurement, Shannon's channel capacity (Nyquist, 1928; Shannon, 1948) imposes an ideal limit on the efficiency of transmission and reception of information. The effect of the "granular nature" of electromagnetic waves on Shannon's channel capacity has been discussed since the advent of the laser (Stern, 1960; Gordon, 1962; Levedev and Levitin, 1963; Takahashi, 1965; She, 1968; Helstrom, 1976; Braginsky and Khalili, 1983). Whereas the quantum-mechanical form of Shannon's channel capacity is very simply stated, its meaning is profound and contains implicitly the rules imposed on the preparation, transmission, and measurement of quantum states. This article reviews the impact of various aspects of quantum measurement on the quantum-mechanical form of Shannon's channel capacity.

Section II discusses the control of the quantum state of electromagnetic waves. A discussion is presented of the characteristics of a coherent state, a squeezed state, a photon number state and an amplitude-squeezed state and the present means of generating them. An ideal laser excited far above threshold can generate a "quasi" coherent state. Optical attenuation can extinguish the difference between the "quasi" coherent state generated by an ideal laser and the "genuine" coherent state.

A squeezed state can be generated via unitary evolution using a phase-conjugate wave. A variety of optical nonlinear processes are candidates (Takahashi, 1965; Yuen and Shapiro, 1979; Walls, 1983; Yurke, 1984). Recently, Slusher *et al.* (1985) observed a squeezed state in a cavity four-wave mixer. An alternative way to generating a squeezed state is the combination of negative feedback via a quantum nondemolition measurement (Yamamoto *et al.*, 1984; Haus and Yamamoto, 1986). The scheme can also generate an amplitude-squeezed state (Yamamoto *et al.* 1986). Regardless of the system's initial state, the quantum nondemolition measurement leaves it in an eigenstate of the measured observable with the measured eigenvalue. The negative feedback is required to keep the system in such an eigenstate continuously, because the system (laser) undergoes unpredictable free motion by coupling to reservoirs (subsystems).

The simultaneous measurement of two conjugate observables inevitably introduces additional noise to resolve the noncommutability of the observables (Haus and Mullen, 1962; Arthurs and Kelly, Jr., 1965; She and Heffner, 1966; Caves, 1982). Types of apparatus for a simultaneous measurement are discussed in Sec. IV.

Optical heterodyning and the beam splitter followed by two single-quadrature measurements are analyzed. The parametric amplifier, the nondegenerate four-wave mixer, the Raman amplifier, the Brillouin amplifier, and the laser amplifier are described as high gain phase insensitive amplifiers. The degenerate parametric amplifier is a phase-sensitive apparatus. The quasi-QND measurement of photon number in an optical Kerr medium is equivalent to a simultaneous measurement of photon number and phase, if the first quasi-QND measurement of photon number is followed by a phase measurement.

Sections V and VI show that a single quadrature measurement and photon counting detection can be done without additional noise. Optical degenerate heterodyning, degenerate parametric amplification, and cavity degenerate four-wave mixing realize such noise-free single quadrature measurements. The photomultiplier and avalanche photodiode are, ideally, noise-free photon counters (Sec. VII).

The quantum-mechanical formulation of Shannon's channel capacity is presented in Sec. VIII, using the negentropy principle of information of Brillouin (1956) and the sampling theorem of Nyquist (1928). The maximum amount of information that can be transmitted by a single photon is infinite at zero temperature and is equal to $\hbar\omega/(\ln 2)kT$ bit at finite temperature. This ultimate information capacity of a photon can be realized only by enormous reduction in channel efficiency, i.e., information rate per bandwidth, however. Thus the quantum noise does not place any limit on the information capacity of a single photon but the thermal noise does. An equivalent but more heuristic statement is that the quantum noise is completely controllable but the thermal noise cannot be fully controlled.

The dependence of the channel capacity on the specific schemes of state preparation and measurement is discussed in Sec. IX. The quantum-mechanical channel capacity derived in Sec. VIII is ideally realized by photon number states and photon counting detection. For the coherent states and squeezed states of large photon number, the simultaneous measurement of two quadrature components and the single quadrature measurement, respectively, are optimum. Although these measurement schemes can recover the highest information for the two states, their channel capacities are smaller than the ultimate limit. The information capacities of a single photon are 1.44 bit for the simultaneous measurement of the coherent state and 2.88 bit for the single quadrature measurement of the squeezed state.

If the coherent state of large photon number is detected by a photon counter, the channel capacity is approximately one-half of the ultimate limit. The halving of the efficiency is due to the fact that equal amounts of information could be carried in the form of photon number and phase, and the photon counting measurement does not utilize the phase information. However, if the coherent state of small photon number is detected by a photon counter, the channel capacity approaches the ultimate limit. This is because the coherent state of small photon number loses its wave nature (phase information).

II. QUANTUM STATES OF ELECTROMAGNETIC WAVES

Information is conveyed electromagnetically by the transmission and reception of electromagnetic wave packets that are quantum states of the electromagnetic field. The transmission involves preparation of these states, the reception is achieved by means of their detection. In this section we review briefly a variety of quantum states of importance, coherent states, in-phase and quadrature-squeezed states, and amplitude-squeezed states and discuss the current means of generating these states.

A. Characteristics and generation of coherent states

A coherent state is the eigenfunction $|\alpha\rangle$ of the photon annihilation operator a:

$$a\,|\alpha\rangle=\alpha\,|\alpha\rangle\ . \tag{2.1}$$

A coherent state is obtained from the vacuum state $|0\rangle$ via the unitary displacement operation $D(\alpha)=\exp(\alpha a^{+}-\alpha^{*}a)$ as follows (Glauber, 1963):

$$|\alpha\rangle=\exp(\alpha a^{+}-\alpha^{*}a)\,|0\rangle\ . \tag{2.2}$$

If the recursion relations for the photon number state $|n\rangle$,

$$a\,|n+1\rangle=\sqrt{n}\,|n\rangle\ , \tag{2.3}$$

$$a^{+}|n\rangle=\sqrt{n+1}\,|n+1\rangle\ , \tag{2.4}$$

are used in (2.2), the expansion coefficients of a coherent state in terms of photon number states are obtained as

$$|\alpha\rangle=e^{-(1/2)|\alpha|^2}\sum_n \frac{\alpha^n}{\sqrt{n!}}\,|n\rangle\ . \tag{2.5}$$

The photon probability distribution $P(n)=|\langle n\,|\,\alpha\rangle|^2$ is Poissonian.

The non-Hermitian operator a can be separated into two Hermitian components a_1 and a_2

$$a=a_1+ia_2 \tag{2.6}$$

which are the "in-phase" and "quadrature" field-component operators.

A coherent state satisfies the following relations:

$$\langle a\rangle=\alpha=\alpha_1+i\alpha_2=\langle a_1\rangle+i\langle a_2\rangle\ , \tag{2.7}$$

$$\langle a^{+}a\rangle=|\alpha|^2=\alpha_1^2+\alpha_2^2=\langle a_1\rangle^2+\langle a_2\rangle^2\ , \tag{2.8}$$

$$\langle\Delta a_1^2\rangle=\langle\Delta a_2^2\rangle=\tfrac{1}{4}\ . \tag{2.9}$$

A coherent state is one of the minimum uncertainty states, which satisfy the Heisenberg uncertainty principle

$$\langle\Delta a_1^2\rangle\langle\Delta a_2^2\rangle=\tfrac{1}{4}\,|\langle[a_1,a_2]\rangle|^2=\tfrac{1}{16}\ . \tag{2.10}$$

Equation (2.9) shows that a coherent state satisfies the minimum uncertainty product (2.10) with equal noises in the in-phase and quadrature components.

Glauber (1951, 1963) has shown that a classical current source generates a coherent state. In a more realistic context, a laser far above threshold generates a sequence of states that are not too far from a coherent state (have little excess noise above the quantum noise). Each state occupies a time interval corresponding to the inverse linewidth of the laser, the phase of the states pertaining to successive time intervals goes through a "random walk." Through attenuation, the excess noise can be made small compared with the quantum noise and thus the attenuated radiation from a laser far above threshold can be made to approach a sequence of coherent states. If the laser is phase locked to a "local oscillator" master laser, the phases of the slave laser follow those of the master oscillator and may be considered controlled (within an uncertainty prescribed by the photon number-phase uncertainty product). Thus one may say that the generation of coherent states for the transmission of information is realizable. The difference in the error distributions $\langle \alpha | \rho | \alpha \rangle$ for a coherent state and the quantum state of an ideal laser is shown in Fig. 1.

B. Characteristics and generation of squeezed states

An ideal squeezed state is another kind of minimum uncertainty state. One quadrature component of the field (say a_1) has smaller fluctuations than the other quadrature component (a_2). Squeezed states appear under several names in the literature: pulsating wave packet (Takahashi, 1965), generalized coherent state (Stoler, 1971), and two-photon coherent state (Yuen, 1976).

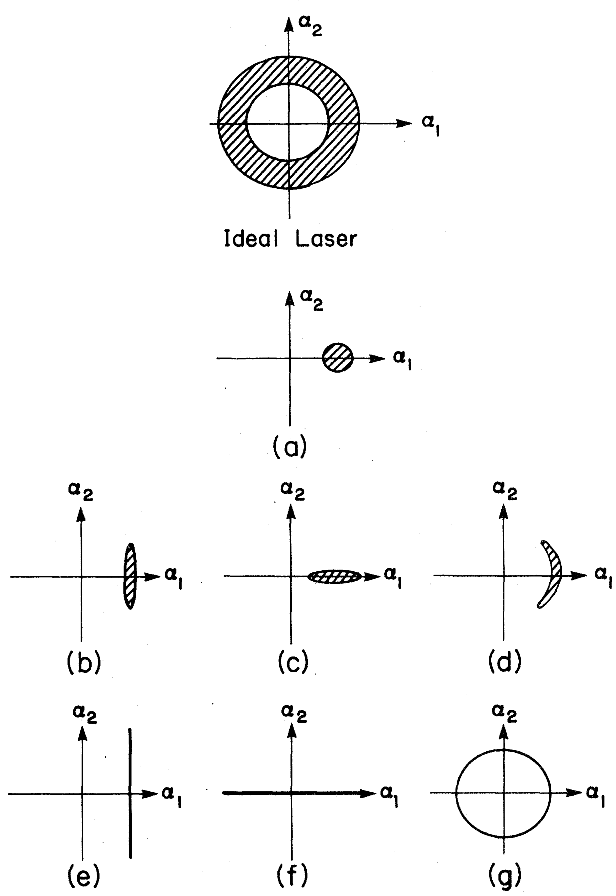

FIG. 1. The error distributions $\langle \alpha | \rho | \alpha \rangle$ in $\alpha_1 - \alpha_2$ space for ideal laser and for a variety of optical quantum states. (a) coherent state, (b) in-phase squeezed state, (c) quadrature-squeezed state, (d) amplitude-squeezed state, (e) quadrature-phase eigenstate, (f) in-phase eigenstate, and (g) photon number eigenstate.

A squeezed state is the eigenfunction $|\mu\nu\alpha\rangle$ of the operator $b = \mu a + \nu a^+$:

$$(\mu a + \nu a^+) |\mu\nu\alpha\rangle = (\mu\alpha + \nu\alpha^*) |\mu\nu\alpha\rangle , \tag{2.11}$$

where

$$|\mu|^2 - |\nu|^2 = 1 , \tag{2.12}$$

so that the commutation relation $[b, b^+] = 1$ is satisfied.

An ideal squeezed state is obtained from the vacuum state $|0\rangle$ by operation with the squeezing operator $S(\zeta) = \exp(\frac{1}{2}\zeta^* a^2 - \frac{1}{2}\zeta a^{+2})$, followed by operation with the displacement operation $D(\alpha)$:

$$|\mu\nu\alpha\rangle = D(\alpha) S(\zeta) |0\rangle , \tag{2.13}$$

where

$$\zeta = re^{i\theta} ,$$

$$|\mu|^2 = \cosh^2 |r| , \tag{2.14}$$

$$|\nu|^2 = \sinh^2 |r| . \tag{2.15}$$

A squeezed state satisfies the following relations:

$$\langle a \rangle = \langle \mu^* b - \nu b^+ \rangle = \alpha = \langle a_1 \rangle + i \langle a_2 \rangle , \tag{2.16}$$

$$\langle a^+ a \rangle = |\alpha|^2 + |\nu|^2 = \langle a_1 \rangle^2 + \langle a_2 \rangle^2 + |\nu|^2 , \tag{2.17}$$

$$\langle \Delta a_1^2 \rangle = \tfrac{1}{4} e^{-2r} , \tag{2.18}$$

$$\langle \Delta a_2^2 \rangle = \tfrac{1}{4} e^{2r} . \tag{2.19}$$

Equations (2.12), (2.18), and (2.19) show that a squeezed state satisfies the minimum uncertainty product and has different amounts of noise in the two quadrature components. Here the complex amplitude axis (α_1, α_2) is rotated by $\theta/2$ so that α_1 and α_2 represent the minor and major axes of the error ellipse. Equations (2.16) and (2.17) indicate that part of the mode energy $|\nu|^2$ is consumed to reduce one of the quadrature noises. Squeezed states are compared with a coherent state in Fig. 1.

An in-phase squeezed state has the reduced quantum noise along the coherent excitation and exhibits sub-Poissonian photon statistics and photon antibunching. A quadrature squeezed state has the reduced quantum noise in quadrature to the coherent excitation and exhibits super-Poissonian photon statistics and photon bunching. The ultimate limit of these states in the "quadrature phase eigenstate" as shown in Fig. 1, which is unrealistic in the sense that it requires an infinite mode energy.

In principle, an ideal squeezed state can be generated by a degenerate parametric amplifier (Takahashi, 1965). The equation for the unitary evolution generating such a state is given by

$$b = \sqrt{G}\, a + \sqrt{G-1}\, a^+ , \tag{2.20}$$

where a is an input mode operator which is assumed to be a coherent state and b is an output operator. The output mode is in a squeezed state.

Degenerate four-wave mixing in an interferometer configuration (Yuen and Shapiro, 1979) and in a cavity configuration (Yurke, 1984) were proposed as other candidates. These schemes reduce to the same basic equation (2.20). Nondegenerate forward four-wave mixing followed by an optical heterodyne detector achieves a relation similar to (2.20) for the intermediate frequency signal (Levenson *et al.*, 1985).

Several experimental efforts are under way toward the generation of squeezed states by these schemes (Bondurant *et al.*, 1984; Levenson, 1985; Slusher *et al.* 1985). Recently Slusher *et al.* (1985) observed the squeezing of vacuum fluctuations in a cavity four-wave mixer.

C. Photon number states, amplitude-squeezed states, and their generation

A photon number state is completely determined by its photon number. The phase is completely random. Photon number states are generated by performing a quantum nondemolition measurement of photon number on a wave packet. The number of photons is unaffected by the measurement and known, after passage of the wave packet through the measurement apparatus. A nonlinear interferometer containing a Kerr medium probed by radiation of (center) frequency different from that of the packet to be measured performs such a QND measurement (Imoto *et al.*, 1985).

An amplitude-squeezed state is a squeezed state that has reduced photon number noise and enhanced phase noise as shown in Fig. 1. Weakly sub-Poissonian photon statistics were observed in resonance fluorescence (Short and Mandel, 1983). It is not necessarily a number-phase minimum uncertainty state.

It has been proposed that an amplitude-squeezed state can be generated by a negative amplitude feedback laser incorporating a quantum nondemolition measurement of photon number (Yamamoto *et al.*, 1984; Yamamoto *et al.*, 1986). The proposed scheme to generate an amplitude-squeezed state is shown in Fig. 2(a). The phase shift of the probe wave in an optical Kerr medium measures the photon number of the laser emission (signal wave) (Imoto, Haus, and Yamamoto, 1985). The photon number fluctuation of the signal wave is measured nondestructively and it is negatively fed back to the laser pump. The photon number fluctuation is reduced, but the phase noise of the signal wave is increased by the phase modulation due to the probe wave intensity noise. As will be shown in Sec. V.E, the back action imposed on the signal phase is $\Delta\varphi = 1/(2\Delta n)$, where Δn is the uncertainty of the photon number measurement. In the limit of large feedback gain, an amplitude-squeezed state which satisfies the minimum uncertainty relation $\Delta n \Delta\varphi = \frac{1}{2}$ can be generated. The same scheme can generate a squeezed state if the QND measurement of a_1 or a_2 (Yurke and Denker, 1984) incorporates negative phase feedback.

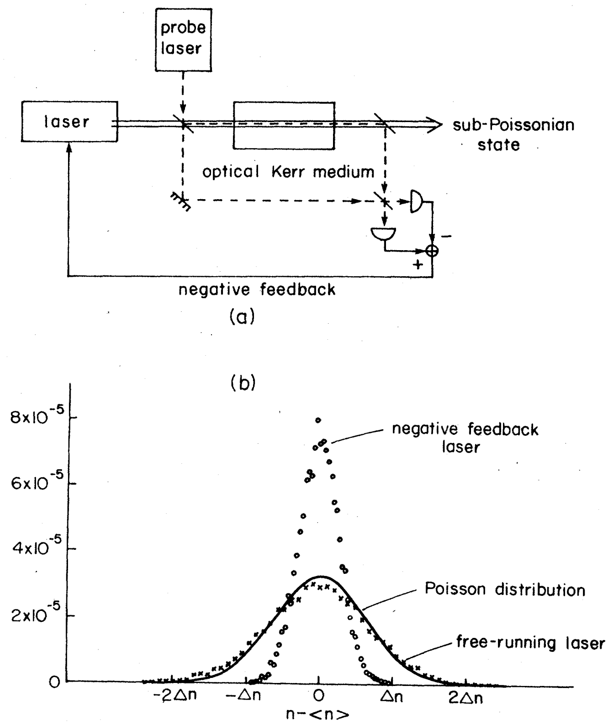

FIG. 2. (a) The configuration of a negative amplitude feedback laser with quantum nondemolition measurement. (b) The observed sub-Poissonian photon statistics in the negative amplitude feedback semiconductor laser using part of the photodetected output.

Sub-Poissonian statistics were actually observed in the negative amplitude feedback semiconductor laser (Machida and Yamamoto, 1986) as shown in Fig. 2(b). In this experiment, however, a conventional destructive photon detector was used instead of the QND measurement of photon number. Under these conditions, amplitude-squeezed states cannot be extracted from the system (Haus and Yamamoto, 1986).

III. GENERAL QUANTUM LIMIT ON THE SIMULTANEOUS MEASUREMENT OF TWO NONCOMMUTING OBSERVABLES

In the preceding section, we have discussed briefly a variety of quantum states of use, or potential use, in the encoding of information. The coherent state has nonzero expectation values for both amplitude and phase, or in-phase and quadrature components a_1 and a_2. Hence both "degrees of freedom" could be used to encode information. In the case when two degrees of freedom represented by two noncommuting quantum observables are used for the transmission of information one has to examine carefully the measurement process. The generation of the encoded signal does not encounter difficulties in principle. One may imagine that the encoding is done at a classical power level, with (almost perfect) control of in-phase and quadrature components, followed by attenuation. The receiver has to detect (measure) these observables simultaneously. This is only possible with the introduction of additional quantum fluctuations (beyond those implied by the Heisenberg uncertainty principle).

In this paper, we use the adjective "simultaneous" measurement to denote the determination of two noncommuting observables of a state, or wave packet. Of course, in a measurement apparatus, the determination of one variable may be delayed with respect to that of the other, and the two variables may not be determined simultaneously. The term double measurement may be more appropriate. However, in communication systems temporal simultaneity is usually required, and thus the use of the term simultaneous seems to be appropriate. Arthurs and Kelly (1965) have studied a simultaneous measurement of two noncommuting variables. They assumed the measurement to be performed by coupling the system, whose observables were to be measured, to two "measurement subsystems." In each of the subsystems one variable is measured, and the system is assumed to be in a minimum uncertainty state. The mean square deviations of the observables measured in this way are, optimally, twice those of the uncertainty principle.

This finding can be summarized as follows. A simultaneous measurement of two noncommuting observables introduces excess noise. The uncertainty product for the measurement is 3 dB larger than the one dictated by the uncertainty principle. The single observable measurement is free from excess noise and, therefore, the uncertainty product for two independent single observable measurements is reduced to the uncertainty principle. Why, then, is such a noise added to the system? We shall show below that additional noise is introduced by the simultaneity of the measurement because two observables measured simultaneously must commute.

Consider a simultaneous "measurement" of two operators a_1 and a_2 achieved by coupling of the original system to two systems described by operators x and y, and by measurement of x and y. The normalized output operators x and y for the measurement of two input operators a_1 and a_2 by means of linear coupling to the system are

$$x=a_1+A \ , \tag{3.1}$$

$$y=a_2+B \ . \tag{3.2}$$

A and B are internal noise operators. An ideal measurement requires

$$\langle x\rangle=\langle a_1\rangle \text{ and therefore } \langle A\rangle=0 \ , \tag{3.3}$$

$$\langle y\rangle=\langle a_2\rangle \text{ and therefore } \langle B\rangle=0 \ . \tag{3.4}$$

Since it is assumed that x and y are simultaneously measured by the two detectors, they must commute:

$$\begin{aligned} 0=[x,y]&=[a_1,a_2]+[a_1,B]+[A,a_2]+[A,B] \\ &=[a_1,a_2]+[A,B] \ , \end{aligned} \tag{3.5}$$

where the third equality uses the fact that the input operators and the internal mode operators are independent and commute with each other

$$[a_1,B]=[A,a_2]=0 \ . \tag{3.6}$$

From (3.5) we conclude that $[B,A]=[a_1,a_2]$. The uncertainty product for A and B is, therefore,

$$\langle \Delta A^2\rangle\langle \Delta B^2\rangle \geq \tfrac{1}{4}\,|\langle [a_1,a_2]\rangle|^2=\tfrac{1}{16} \ . \tag{3.7}$$

If the input operators and internal-mode operators are uncorrelated, the uncertainty product for x and y is given by

$$\begin{aligned} \langle \Delta x^2\rangle\langle \Delta y^2\rangle &=(\langle \Delta a_1^2\rangle+\langle \Delta A^2\rangle)(\langle \Delta a_2^2\rangle+\langle \Delta B^2\rangle) \\ &\geq \langle \Delta a_1^2\rangle\langle \Delta a_2^2\rangle+\langle \Delta A^2\rangle\langle \Delta B^2\rangle \\ &\quad +2(\langle \Delta a_1^2\rangle\langle \Delta a_2^2\rangle\langle \Delta A^2\rangle\langle \Delta B^2\rangle)^{1/2} \\ &\geq 4\langle \Delta a_1^2\rangle\langle \Delta a_2^2\rangle=\tfrac{1}{4} \ . \end{aligned} \tag{3.8}$$

Here the equality holds when $\langle \Delta A^2\rangle=\langle \Delta a_1^2\rangle$ and $\langle \Delta B^2\rangle=\langle \Delta a_2^2\rangle$.

The above discussion shows that a simultaneous measurement requires internal-mode fluctuations to allow commutation of the two simultaneously measured output operators x and y. These fluctuations in turn increase the uncertainty of the measurement by at least 3 dB from the uncertainty product of the input operators. If one lifts the assumption that the input operators and internal-mode operators are uncorrelated, it is possible to obtain $\langle \Delta x^2\rangle\langle \Delta y^2\rangle=0$. The additional noise encountered in the simultaneous measurement of two noncommuting observables has various origins depending upon the specific measurement apparatus employed. Detectors measuring two noncommuting observables fall into two general types.

One type incorporates an amplifier of high gain. After amplification of the signal to a classical power level, the signal can be measured with no uncertainty; a simultaneous measurement can be performed on both the in-phase and quadrature components, or on the amplitude or phase. The uncertainty principle is obeyed by virtue of the fact that noise was introduced in the amplification process.

The other type involves directly simultaneous measurements on the two noncommuting variables with no preamplification. An example is the beam splitter followed by measurements of the in-phase component on one output port of the beam splitter, of the quadrature component at the other output port. A quantum nondemolition measurement of photon number, followed by a phase measurement is another example. Heterodyne detection resembles the high gain amplifier system followed by a detector. The analysis of heterodyne detection always assumes effectively a high gain so that the currents produced in the photodetector can be treated classically. In effect, every measurement of a quantum observable is performed by instruments that operate (at their output) in a classical environment, as pointed out by Bohr.

An amplifier with different gains G_1 and G_2, for the in-phase component a_1 and quadrature component a_2, respectively, can be viewed as either a component of a simultaneous measurement of two noncommuting observables if $G_1 \gg 1$ and $G_2 \gg 1$, or as a component of a single observable measurement, if one of the gains is close to unity or less than unity. It is an easy matter to set up the general formalism for this amplifier type which then serves as a means of comparison for all measurement apparati.

IV. THEORY OF A LINEAR, PHASE-SENSITIVE AMPLIFIER

In this section we write down the general theory of a linear, phase-sensitive amplifier with gain G_1 for the (power of the) in-phase component a_1 and G_2 for the (power of the) quadrature component a_2. The noise introduced by the amplifier must account for the quantum-mechanical uncertainty of a simultaneous measurement (if $G_1 \gg 1$ and $G_2 \gg 1$).

The two quadrature components of the amplified output operators b_1 and b_2, in terms of the input operators a_1 and a_2 are given by (Haus and Mullen, 1962; Caves, 1982)

$$b_1 = \sqrt{G_1}a_1 + F_1 , \tag{4.1}$$

$$b_2 = \sqrt{G_2}a_2 + F_2 . \tag{4.2}$$

Here, F_1 and F_2 are the (internal-mode) fluctuation operators. The mode operators b_1, b_2 and a_1, a_2 must both satisfy the boson commutation relations

$$[a_1, a_2] = [b_1, b_2] = \frac{i}{2} , \tag{4.3}$$

because they are both operator (excitation) amplitudes of the same kind of mode (i.e., a wave packet "fed" into the input "transmission line" emerges as an amplified wave packet in the output "transmission line"). In order that (4.3) be obeyed, one finds

$$[F_1, F_2] = \frac{i}{2}(1 - \sqrt{G_1 G_2}) . \tag{4.4}$$

Here, the assumption used is that the input operators a_1 and a_2 are independent of the internal fluctuation operators, $[a_1, F_2] = [a_2, F_1] = 0$. One may define the normalized output signal by

$$a_{1,\text{eff}} \equiv \frac{b_1}{\sqrt{G_1}} , \quad a_{2,\text{eff}} \equiv \frac{b_2}{\sqrt{G_2}} . \tag{4.5}$$

The mean square signal output is then

$$\overline{\langle a_{1,\text{eff}} \rangle^2} = \frac{\overline{\langle b_1 \rangle^2}}{G_1} , \quad \overline{\langle a_{2,\text{eff}} \rangle^2} = \frac{\overline{\langle b_2 \rangle^2}}{G_2} , \tag{4.6}$$

and the noise output is, in the two components

$$\langle \Delta a_{1,\text{eff}}^2 \rangle = \langle \Delta a_1^2 \rangle + \frac{\langle \Delta F_1^2 \rangle}{G_1} , \tag{4.7}$$

$$\langle \Delta a_{2,\text{eff}}^2 \rangle = \langle \Delta a_2^2 \rangle + \frac{\langle \Delta F_2^2 \rangle}{G_2} . \tag{4.8}$$

Here it is assumed again that a and F are uncorrelated.

The high gain phase-sensitive amplifier with $G_1 \gg 1$ and $G_2 \gg 1$ is a special case of the measurement of two noncommuting observables as discussed in the preceding section. The normalized output operators commute when the gains G_1 and G_2 are high, and can be identified with the observables x and y of the preceding section. Indeed, using (4.4)

$$\begin{aligned}[a_{1,\text{eff}}, a_{2,\text{eff}}] &= [a_1, a_2] + \frac{[F_1, F_2]}{\sqrt{G_1 G_2}} \\ &= \frac{i}{2} + \frac{i}{2}\left[\frac{1}{\sqrt{G_1 G_2}} - 1\right] \\ &= \frac{i}{2}\frac{1}{\sqrt{G_1 G_2}} .\end{aligned} \tag{4.9}$$

Thus, when $\sqrt{G_1 G_2} \gg 1$, then the two observables commute and can be measured independently. We can see from the preceding derivation that conservation of commutation relations of the operators led to commutability of the measured observables, because the observables are the output operators divided by the square roots of the gains.

Next, we determine the uncertainty product of $\Delta a_{1,\text{eff}}$ and $\Delta a_{2,\text{eff}}$. Using (4.4), we obtain

$$\begin{aligned}\frac{\langle \Delta F_1^2 \rangle \langle \Delta F_2^2 \rangle}{G_1 G_2} &\geq \frac{1}{4} \frac{|\langle [F_1, F_2] \rangle|^2}{(G_1 G_2)} \\ &= \frac{1}{16}\left[1 - \frac{1}{\sqrt{G_1 G_2}}\right]^2\end{aligned} \tag{4.10}$$

and from (4.1) and (4.2)

$$\begin{aligned}\langle \Delta a_{1,\text{eff}}^2 \rangle \langle \Delta a_{2,\text{eff}}^2 \rangle &= \left[\langle \Delta a_1^2 \rangle + \frac{\langle \Delta F_1^2 \rangle}{G_1}\right]\left[\langle \Delta a_2^2 \rangle + \frac{\langle \Delta F_2^2 \rangle}{G_2}\right] \\ &\geq \langle \Delta a_1^2 \rangle \langle \Delta a_2^2 \rangle + \frac{\langle \Delta F_1^2 \rangle}{G_1}\frac{\langle \Delta F_2^2 \rangle}{G_2} + 2\left[\langle \Delta a_1^2 \rangle \langle \Delta a_2^2 \rangle \frac{\langle \Delta F_1^2 \rangle}{G_1}\frac{\langle \Delta F_2^2 \rangle}{G_2}\right]^{1/2} \\ &\geq \frac{1}{16} + \frac{1}{16}\left[1 - \frac{1}{\sqrt{G_1 G_2}}\right]^2 + 2\frac{1}{16}\left[1 - \frac{1}{\sqrt{G_1 G_2}}\right] \geq \frac{1}{16}\left[2 - \frac{1}{\sqrt{G_1 G_2}}\right]^2 .\end{aligned} \tag{4.11}$$

The equality holds when the input is a minimum uncertainty state, and when

$$\frac{\langle \Delta F_1^2 \rangle}{G_1} = \langle \Delta a_1^2 \rangle \tag{4.12}$$

and

$$\frac{\langle \Delta F_2^2 \rangle}{G_2} = \langle \Delta a_2^2 \rangle \ . \tag{4.13}$$

Equation (4.11) is similar to (3.8), but with a difference. If the product of gain coefficients $G_1 G_2$ does not satisfy the condition of $G_1 G_2 \gg 1$, then the uncertainty product can be smaller than $\frac{1}{4}$. This seems to be in contradiction with the general quantum limit for a simultaneous measurement discussed in the preceding section. The reason for this discrepancy is that a classical measurement free of additive noise can be performed only when $G_1 G_2 \gg 1$. Indeed, according to (4.9) $a_{1,\mathrm{eff}}$ and $a_{2,\mathrm{eff}}$ do not commute and, therefore, cannot be measured simultaneously, when $G_1 G_2$ is not much greater than unity.

The gain factors $\sqrt{G_1}$ and $\sqrt{G_2}$ suggest a classification of the amplifiers studied in this section. If $\sqrt{G_1}$ and $\sqrt{G_2}$ are independent of a_1 and a_2, the amplifier is linear and belongs to one of three categories (Caves, 1982):

$\sqrt{G_1} = \sqrt{G_2}$ (phase-preserving amplifier) ,

$\sqrt{G_1} = -\sqrt{G_2}$ (phase-conjugate amplifier) ,

$|G_1| \neq |G_2|$ (phase-sensitive amplifier) ,

In the special case of $G_1 G_2 = 1$, the amplifier does not add noise [see (4.4)] and the uncertainty product of (4.11) is reduced to $\frac{1}{16}$. The product $G_1 G_2$ can be kept equal to unity by making $G_1 = 1/G_2$ and $G_1 \geq 1$. In this case a_1 can be measured with no additive noise and the information on a_2 is sacrificed. Takahashi (1965) was first to discuss a degenerate parametric amplifier as such an amplifier. Caves (1982) discussed it again recently.

There is another mode for an ideal single observable measurement (Yurke and Denker, 1984). It is

$$\frac{\langle \Delta F_1^2 \rangle}{G_1} \ll \langle \Delta a_1^2 \rangle \ , \quad \frac{\langle \Delta F_2^2 \rangle}{G_2} \gg \langle \Delta a_2^2 \rangle \ . \tag{4.14}$$

The measurement of a_1 is ideal, but the information on a_2 is lost. A parametric amplifier with a zero-mean squeezed state fed into the idler channel will be treated as an example later on. Figure 3 summarizes the dependences on gain of the internal fluctuations.

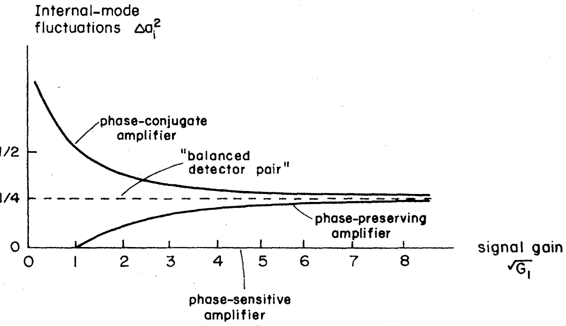

FIG. 3. Internal-mode fluctuations vs signal gain for a phase preserving amplifier, a phase-conjugate amplifier, a phase-sensitive amplifier, and "balanced detector pair."

The linear, phase-sensitive amplifier with $G_1 \gg 1$ and $G_2 \gg 1$ is a general example of a measurement apparatus that measures two noncommuting observables simultaneously. Specific examples are the laser amplifier, parametric amplifier, four-wave mixer, etc. In the next section we look in more detail at "devices" for a simultaneous measurement and compare them with the results of this section.

V. "DEVICES" FOR SIMULTANEOUS MEASUREMENT OF TWO NONCOMMUTING OBSERVABLES

A. The ideal laser amplifier

An ideal laser amplifier, with gain produced by a perfectly inverted medium, can be described by equations analogous to (4.1) with $G_1 = G_2$. It is a linear, phase-preserving amplifier.

The output operator $b = b_1 + i b_2$ may be written in terms of the input operator $a = a_1 + i a_2$,

$$b = \sqrt{G} a + F \ , \tag{5.1}$$

where $F = F_1 + i F_2$ is the noise operator. In order to preserve the commutation relations

$$[b, b^+] = [a, a^+] = 1 \ , \tag{5.2}$$

one must have

$$[F, F^+] = 1 - G \ . \tag{5.3}$$

When $G < 1$ (attenuation), F denotes a zero-point fluctuation added by a "loss oscillator." For a gain medium, $G > 1$, F and F^+ change their roles as creation and annihilation operators, and F^+ denotes a zero-point fluctuation annihilation operator. This relation can be interpreted in another way. The expectation value for $b^+ b$ with a coherent state $|\alpha\rangle$ into the signal channel, $a|\alpha\rangle = \alpha|\alpha\rangle$, and a vacuum state for the noise, $\langle F F^+ \rangle = 0$, is

$$\langle \alpha | b^+ b | \alpha \rangle = (G-1)(|\alpha|^2 + 1) + |\alpha|^2 \ . \tag{5.4}$$

For every induced signal photon there is added one spontaneously emitted noise photon. This reasoning was used in the early days of the maser to predict maser noise performance (Shimoda, Takahashi, and Townes, 1957; Strandberg, 1963). An injection-locked laser oscillator has the same quantum limit, even though the signal gains are different for in-phase and quadrature-phase components (Haus and Yamamoto, 1984).

B. Parametric amplifier and four-wave mixer

The input to a parametric amplifier and a four-wave mixer consists of two modes. A signal wave at ω_s and an idler wave at ω_i, that are coupled with each other by second- and third-order nonlinear processes produced by the intense pump wave at ω_p. The basic configuration is shown in Fig. 4. Energy conservation requires

$$\omega_s+\omega_i=\omega_p \quad \text{parametric amplifier} , \tag{5.5}$$

$$\omega_s+\omega_i=2\omega_p \quad \text{four-wave mixer} . \tag{5.6}$$

If the pump wave is in a coherent state and its intensity is sufficiently high, it can be treated as a classical wave. The evolution equations for the signal and idler waves are (Yariv and Louisell, 1966)

$$b_s=\sqrt{G}\,a_s+\sqrt{G-1}\,a_i^+ , \tag{5.7}$$

$$b_i=\sqrt{G-1}\cdot a_s^+ +\sqrt{G}\,a_i , \tag{5.8}$$

where $a_s(b_s)$ and $a_i(b_i)$ are input (output) operators for the signal and idler waves, and $\sqrt{G}$ is the signal gain.

Suppose the signal channel is taken as the output. The device is a phase-preserving linear amplifier with signal gain of $\sqrt{G_1}=\sqrt{G_2}=\sqrt{G}$. The mean and variance of a normalized output $y_s\equiv b_s/\sqrt{G}$ are

$$\langle y_{s1}\rangle=\langle a_{s1}\rangle , \tag{5.9}$$

$$\langle \Delta y_{s1}^2\rangle=\langle \Delta a_{s1}^2\rangle+\left[1-\frac{1}{G}\right]\langle \Delta a_{i1}^2\rangle , \tag{5.10}$$

$$\langle y_{s2}\rangle=\langle a_{s2}\rangle , \tag{5.11}$$

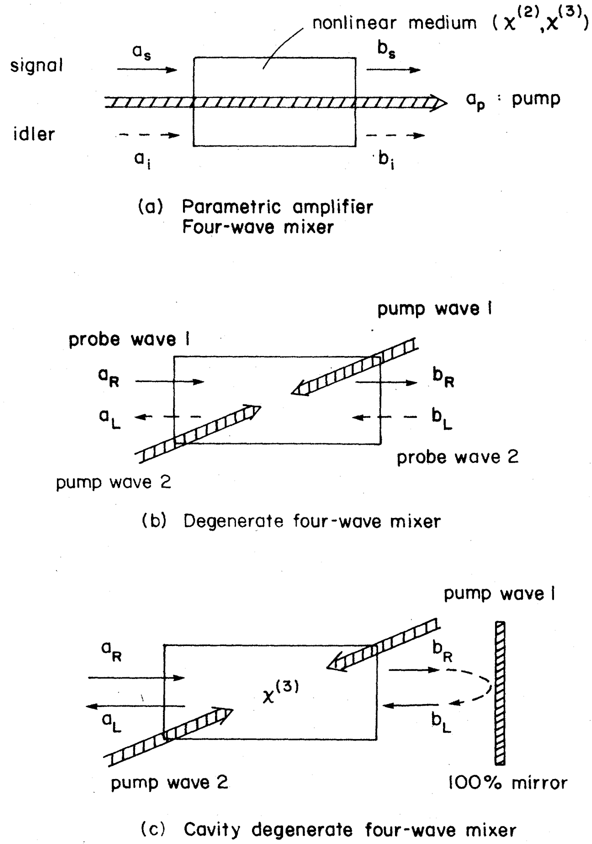

FIG. 4. Basic configurations of (a) a parametric amplifier, (b) a degenerate four-wave mixer, and (c) a cavity degenerate four-wave mixer.

$$\langle \Delta y_{s2}^2\rangle=\langle \Delta a_{s2}^2\rangle+\left[1-\frac{1}{G}\right]\langle \Delta a_{i2}^2\rangle . \tag{5.12}$$

A simultaneous measurement of a_{s1} and a_{s2} can be achieved for a coherent state input when a_i is in a conventional vacuum state $|0\rangle$. If the input to the idler is a zero-mean squeezed state such that $\langle \Delta a_{i1}^2\rangle \ll \langle \Delta a_{s1}^2\rangle$, then no noise is added by the internal fluctuations. Of course, a measurement of the conjugate variable becomes impossible because the noise $\langle \Delta y_s^2\rangle$ goes to infinity. The measurement "degenerates" into that of a single observable with no additional noise.

Suppose the idler channel is taken as the output. Then the device is a phase-conjugate linear amplifier with signal gain of $\sqrt{G_1}=-\sqrt{G_2}=\sqrt{G-1}$. The mean and variance of a normalized output $y_i\equiv b_i/\sqrt{G-1}$ are

$$\langle y_{i1}\rangle=\langle a_{s1}\rangle , \tag{5.13}$$

$$\langle \Delta y_{i1}^2\rangle=\langle \Delta a_{s1}^2\rangle+\left[1+\frac{1}{G-1}\right]\langle \Delta a_{i1}^2\rangle , \tag{5.14}$$

$$\langle y_{i2}\rangle=-\langle a_{s2}\rangle , \tag{5.15}$$

$$\langle \Delta y_{i2}^2\rangle=\langle \Delta a_{s2}^2\rangle+\left[1+\frac{1}{G-1}\right]\langle \Delta_{i2}^2\rangle . \tag{5.16}$$

The result is the same as that of the linear phase-preserving amplifier in the limit of high signal gain $G\gg 1$. However, there is a subtle difference. As shown in Fig. 3, a phase-conjugate linear amplifier is more noisy than a phase-preserving linear amplifier unless the signal gain is high.

A degenerate four-wave mixer is shown in Fig. 4(b), where the signal, idler, and pump waves are at the same frequency. There are still two input modes a_R and b_L. The evolution equations are (Yuen and Shapiro, 1979)

$$b_R=\sqrt{G}\,a_R+e^{i\varphi}\sqrt{G-1}\,b_L^+ , \tag{5.17}$$

$$a_L=\sqrt{G}\,b_L+e^{i\varphi}\sqrt{G-1}\,a_R^+ , \tag{5.18}$$

Since (5.17) and (5.18) are of the same form as (5.7) and (5.8), the degenerate four-wave mixer performs like the nondegenerate parametrix amplifier.

The above result can be applied to any optical amplifier with a parametric interaction process among several bosonic modes, such as the Raman amplifier and the Brillouin amplifier. The internal-mode fluctuations are always added to the signal during such an amplification process. They are the zero-point fluctuations of a lattice vibrational mode (optical phonon) in the Raman amplifier, and the zero-point fluctuations of an acoustic phonon mode in the Brillouin amplifier (Louisell, Yariv, and Siegman, 1961). The effects of quantum noise of a pump wave and of an internal loss are not included in the above discussion, because they are not fundamental limiting factors.

C. Heterodyning

Yuen and Shapiro (1980) have developed the detailed theory of heterodyning with a photon detector, including the cases of heterodyning with squeezed states as the input in either the signal channel, idler channel, or both. The schematic is shown in Fig. 5. The beam splitter combines the incoming signal wave with the local oscillator wave. The local oscillator is assumed to be much more intense than the incoming signal. The photon detector has quantum efficiency η.

The analysis of Yuen and Shapiro can be summarized in a set of simple equations that give the expectation value and the mean square fluctuations of the current in the signal channel. The current variable y is so normalized that its expectation value gives the number of photons detected in one observation time interval. The current has an in-phase (cosine) component y_c, and a quadrature (sine) component y_s. The expectation values are

$$\langle y_c \rangle = \langle a_{s1} \rangle , \tag{5.19}$$

$$\langle y_s \rangle = \langle a_{s2} \rangle , \tag{5.20}$$

and the mean square fluctuations are

$$\langle \Delta y_c^2 \rangle = \langle \Delta a_{s1}^2 \rangle + \langle \Delta a_{i1}^2 \rangle + \frac{1-\varepsilon}{\varepsilon}(\langle \Delta b_{i1}^2 \rangle + \langle \Delta b_{s1}^2 \rangle) + \frac{1-\eta}{2\eta\varepsilon} . \tag{5.21}$$

The terms in (5.21) represent the in-phase noise of the signal wave, the zero-point fluctuations of the vacuum modes in the image band of the signal arm, the zero-point fluctuations in the signal and image bands of the local oscillator arm, and the zero-point fluctuations introduced by the finite quantum efficiency of the photodetector, respectively. An analogous relation holds for the quadrature fluctuations $\langle \Delta y_s^2 \rangle$ with subscripts 1 changed to 2. Here the beating noise between the coherent excitation of the signal wave and the zero-point fluctuations (Shapiro, 1985) is neglected.

In the limit of $\eta \rightarrow 1$ and $\varepsilon \rightarrow 1$, but still preserving the local oscillator gain much larger than unity, Eq. (5.21) and the analogous expression for $\langle \Delta y_s^2 \rangle$ reduce to

$$\langle \Delta y_c^2 \rangle = \langle \Delta a_{s1}^2 \rangle + \langle \Delta a_{i1}^2 \rangle , \tag{5.22}$$

$$\langle \Delta y_s^2 \rangle = \langle \Delta a_{s2}^2 \rangle + \langle \Delta a_{i2}^2 \rangle . \tag{5.23}$$

Equations (5.19), (5.20), (5.22), and (5.23) are identical with (5.9)—(5.12), in the limit when the gain in the latter goes to infinity, $G \rightarrow \infty$. Thus the heterodyne receiver performs like a parametric amplifier, followed by a noise-free detector (a practical detector can be considered noise free if preceded by an amplifier of high gain).

Equations (5.22) and (5.23) suggest that heterodyning is an abstract quantum measurement of $a_s + a_i^+$ and that the in-phase and quadrature parts of y commute and can be measured simultaneously. It has been argued, however, that there is an uncertainty product of order ω_{if}/ω_0 for the simultaneous measurement of y_c and y_s (Caves, 1981).

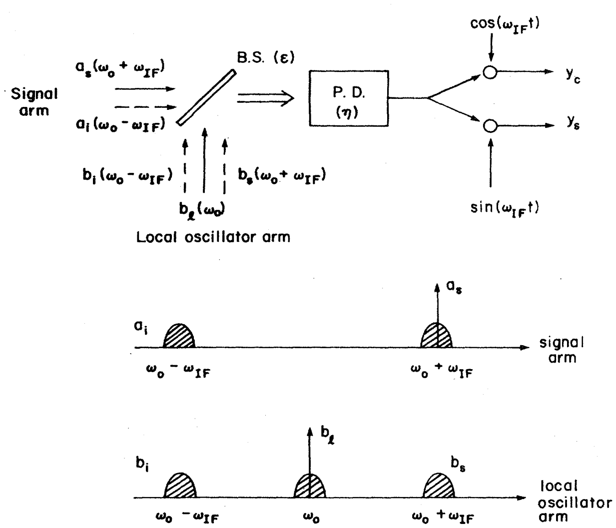

FIG. 5. Basic configuration of an optical heterodyne receiver. There are two modes in the signal arm, an input signal a_s, and a vacuum mode a_i in image band $\omega_0 - \omega_{IF}$. There are three modes in the local oscillator arm, a local oscillator wave b_l and two vacuum modes b_i and b_s.

D. The beam splitter

The beam splitter followed by two independent single quadrature measurements does not incorporate any gain before detection. (As pointed out earlier, the detection process reduces the output to a classical observable whose determination is not subject to uncertainty.) It differs in this respect from the amplifier-detector apparatus and its noise has a different origin from that of the amplification process.

A beam splitter followed by two independent single quadrature measurements is shown in Fig. 6. Practical detectors for single quadrature measurements will be discussed in the next section. It might be optical degenerate heterodyning (homodyning), degenerate parametric amplification, or cavity degenerate four-wave mixing. All of

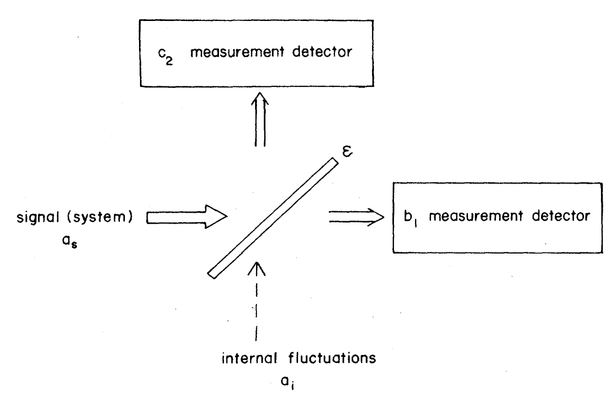

FIG. 6. A beam splitter followed by two independent single quadrature measurements.

these are free from excess noise. Oliver (1962) was first to consider this type of simultaneous measurement, in which he assumed two optical homodyne receivers each performing a single quadrature measurement. Unfortunately he obtained the incorrect conclusion that the Heisenberg uncertainty principle can be realized by such a detector by erroneously dropping a factor of 4 in the noise calculation for the photon number. The operators for the two output arms of the beam splitter are

$$b=\varepsilon^{1/2}a_s+(1-\varepsilon)^{1/2}a_i \;, \tag{5.24}$$

$$c=-(1-\varepsilon)^{1/2}a_s+\varepsilon^{1/2}a_i \;, \tag{5.25}$$

where a_i is the vacuum mode incident on the beam splitter from the open port which is indicated by "internal fluctuations" in Fig. 6.

Suppose that a_{s1} is inferred from a measurement in the output arm (b) of the beam splitter and a_{s2} is inferred from the output arm (c). Then, ε should be chosen to be $\frac{1}{2}$ to get equal accuracies in the two measurements. If one defines the variables to be measured by

$$a_{1,\mathrm{eff}}=\frac{b_1}{\sqrt{\varepsilon}}=\sqrt{2}b_1 \tag{5.26}$$

and

$$a_{2,\mathrm{eff}}=\frac{c_2}{\sqrt{1-\varepsilon}}=\sqrt{2}c_2 \;, \tag{5.27}$$

one obtains $[a_{1,\mathrm{eff}},a_{2,\mathrm{eff}}]=-[a_{s1},a_{s2}]+[a_{i1},a_{i2}]=0$. This is the same relation as (3.5). Furthermore,

$$\langle a_{1,\mathrm{eff}}\rangle=\langle a_{s1}\rangle \;, \tag{5.28}$$

$$\langle a_{2,\mathrm{eff}}\rangle=\langle a_{s2}\rangle \;, \tag{5.29}$$

and

$$\langle \Delta a^2_{1,\mathrm{eff}}\rangle=\langle \Delta a^2_{s1}\rangle+\langle \Delta a^2_{i1}\rangle \;, \tag{5.30}$$

$$\langle \Delta a^2_{2,\mathrm{eff}}\rangle=\langle \Delta a^2_{s2}\rangle+\langle \Delta a^2_{i2}\rangle \;. \tag{5.31}$$

This is of the same form as the relations for the parametric amplifier of high gain. Here the additional noise contributions arise from the coupling to the zero-point fluctuations of the "unused" input port of the beam splitter.

E. Quasiquantum nondemolition measurement followed by phase measurement

One possible configuration of the quantum nondemolition measurement of photon number is shown in Fig. 2(a) (Imoto, Haus, and Yamamoto, 1985). A different configuration based on similar principles was discussed by Braginsky and Vyatchanin (1981, 1982). A signal wave propagates along an optical Kerr medium without suffering any loss. The refractive index of the Kerr medium, changed by the signal wave intensity, is probed as a phase shift of the probe wave passing through the Kerr medium. Thus the photon number of the signal wave can be measured nondestructively. The phase shift φ_p of the probe wave produced by the signal wave is proportional to the signal photon number N_s:

$$\varphi_p=\sqrt{F}N_s \;, \tag{5.32}$$

where F is a constant proportional to the third-order nonlinear coefficient $\chi^{(3)}$, the interaction length L and the signal and probe frequencies. Any phase shift produced on the probe wave by the probe photon number fluctuation can be either canceled by passing the probe wave through a Kerr medium with opposite sign of $\chi^{(3)}$, or can be avoided through proper use of resonant excitations. The expectation value of N_s is

$$\langle N_s\rangle=\frac{1}{\sqrt{F}}\langle \varphi_p\rangle \;. \tag{5.33}$$

The measurement accuracy of N_s is limited by the natural fluctuations of φ_p to

$$\langle \Delta N_s^2\rangle_{\mathrm{meas}}\geq\frac{1}{F}\langle \Delta\varphi_p^2\rangle \;. \tag{5.34}$$

If $\langle \Delta\varphi_p^2\rangle$ is made very small, then the accuracy of the measurement is correspondingly increased. In the limit of $\langle \Delta\varphi_p^2\rangle/F\to 0$, this measurement scheme can be considered as the quantum nondemolition measurement of photon number. Repeated measurements performed on the same wave packet at different positions and times give the same value of photon number. The probe acts on the signal wave by its own photon number fluctuations. A relation of the form (5.32) holds with subscripts s and p interchanged. Therefore, the measurement introduces a signal-phase perturbation

$$\langle \Delta\varphi_s^2\rangle_{\mathrm{meas}}=\langle \Delta N_p^2\rangle F \;. \tag{5.35}$$

If the probe wave is in a minimum uncertainty (coherent) state we obtain

$$\langle \Delta N_p^2\rangle\langle \Delta\varphi_p^2\rangle=\tfrac{1}{4} \;. \tag{5.36}$$

Thus,

$$\langle \Delta N_s^2\rangle_{\mathrm{meas}}\langle \Delta\varphi_s^2\rangle_{\mathrm{meas}}\geq\tfrac{1}{4} \;. \tag{5.37}$$

The phase perturbation produced by the measurement and the uncertainty in the measurement of photon number obey the minimum uncertainty product.

The quasi-QND measurement of photon number with a finite measurement accuracy can be part of a measurement of two noncommuting variables if the wave emerging from the QND measurement apparatus is subjected to a phase measurement. The signal enters the QND measurement apparatus with fluctuations $\langle \Delta N_s^2\rangle_{\mathrm{in}}$ and $\langle \Delta\varphi_s^2\rangle_{\mathrm{in}}$. The fluctuation at the output of the apparatus are

$$\langle \Delta N_s^2\rangle_{\mathrm{tot}}=\langle \Delta N_s^2\rangle_{\mathrm{in}}+\langle \Delta N_s^2\rangle_{\mathrm{meas}} \;, \tag{5.38}$$

$$\langle \Delta\varphi_s^2\rangle_{\mathrm{tot}}=\langle \Delta\varphi_s^2\rangle_{\mathrm{in}}+\langle \Delta\varphi_s^2\rangle_{\mathrm{meas}} \;, \tag{5.39}$$

their uncertainty product is

$$\langle \Delta N_s^2\rangle_{\mathrm{tot}}\langle \Delta\varphi_s^2\rangle_{\mathrm{tot}}\geq\tfrac{1}{2}+\langle \Delta N_s^2\rangle_{\mathrm{meas}}\langle \Delta\varphi_s^2\rangle_{\mathrm{in}}+\langle \Delta N_s^2\rangle_{\mathrm{in}}\langle \Delta\varphi_s^2\rangle_{\mathrm{meas}} \;. \tag{5.40}$$

The equality sign holds when the incoming signal is in a minimum uncertainty state,

$$\langle \Delta N_s^2 \rangle_{\rm in} \langle \Delta \varphi_s^2 \rangle_{\rm in} = \tfrac{1}{4} \ . \qquad (5.41)$$

When we introduce (5.37) and further use (5.41) we find that $\langle \Delta N_s^2 \rangle_{\rm tot} \langle \Delta \varphi_s^2 \rangle_{\rm tot}$ reaches a minimum of 1 when the adjustment is made $\langle \Delta \varphi_s^2 \rangle_{\rm in} = \langle \Delta \varphi_s^2 \rangle_{\rm meas}$. This shows that the quasi-QND measurement of photon number, followed by a phase measurement, under proper adjustment of the measurement conditions leads to a doubling of the noise associated with the in-phase and quadrature components, or photon number and phase.

VI. QUANTUM NOISE OF SINGLE QUADRATURE MEASUREMENT DETECTORS

There are three single quadrature measurement detectors: optical degenerate heterodyne detection, the degenerate parametric amplifier, and the cavity degenerate four-wave mixer.

A. Degenerate heterodyning (homodyning)

Degenerate heterodyne detection measures only one component of the signal wave (Haus and Townes, 1962; Oliver, 1962). The operation of the detector is contained in that of the heterodyne detector with signal and idler channels merged. This means that the noise in the signal channel and the idler channel are one and the same and must be counted only once. Instead of (5.22) and (5.23), one has

$$\langle \Delta y_1^2 \rangle = \langle \Delta a_{s1}^2 \rangle \qquad (6.1)$$

for an in-phase detector, and

$$\langle \Delta y_2^2 \rangle = \langle \Delta a_{s2}^2 \rangle \qquad (6.2)$$

for the quadrature detector. Of course,

$$\langle y_1 \rangle = \langle a_{s1} \rangle \qquad (6.3)$$

for the in-phase detector, and

$$\langle y_2 \rangle = \langle a_{s2} \rangle \qquad (6.4)$$

for the quadrature detector.

These relations determine the signal-to-noise ratio, which in turn fixes the channel capacity achievable with degenerate heterodyne detection.

B. Degenerate parametric amplifier

A degenerate parametric amplifier is a special case of a parametric amplifier in which the signal frequency is equal to half of the pump frequency, $\omega_s = \frac{1}{2}\omega_p$. The evolution equations for the signal and idler waves, (5.7) and (5.8), are reduced to the following single equation:

$$b_s = \sqrt{G} a_s + \sqrt{G-1} a_s^{+} = e^{r} a_{s1} + i e^{-r} a_{s2} \ . \qquad (6.5)$$

The signal gain for the a_{s1} component is $\sqrt{G_1} = e^{r} = \sqrt{G} + \sqrt{G-1}$ and the signal gain for the a_{s2} component is $\sqrt{G_2} = e^{-r} = \sqrt{G} - \sqrt{G-1}$, which corresponds to loss. The normalized output $y_1 \equiv b_{s1}/\sqrt{G_1}$ and $y_s \equiv b_{s2}/\sqrt{G_2}$ satisfy

$$\langle y_1 \rangle = \langle a_{s1} \rangle \ , \qquad (6.6)$$

$$\langle \Delta y_1^2 \rangle = \langle \Delta a_{s1}^2 \rangle \ , \qquad (6.7)$$

$$\langle y_2 \rangle = \langle a_{s2} \rangle \ , \qquad (6.8)$$

$$\langle \Delta y_2^2 \rangle = \langle \Delta a_{s2}^2 \rangle \ . \qquad (6.9)$$

Equations (6.7) and (6.9) suggest that a degenerate parametric amplifier measures a_{s1} and a_{s2} without excess noise. As discussed in Sec. V, the internal-mode fluctuations vanish if $G_1 G_2$ is equal to unity. A degenerate parametric amplifier satisfies this condition. However, the quadrature component, a_{s2} is actually attenuated and cannot be measured. Therefore the performance of the degenerate parametric amplifier at high gain is the same as that of degenerate optical heterodyning.

The output signal b_s of a degenerate parametric amplifier is in a squeezed state with quadrature noise magnitudes $\langle \Delta b_{s1}^2 \rangle = e^{2r}/4$ and $\langle \Delta b_{s2}^2 \rangle = e^{-2r}/4$ that satisfy the minimum uncertainty product of $\langle \Delta b_{s1}^2 \rangle \langle \Delta b_{s2}^2 \rangle = \frac{1}{16}$.

C. Cavity degenerate four-wave mixer

Suppose the unused port of the degenerate four-wave mixer, b_L in Fig. 4(b), is closed by a perfectly reflecting mirror as shown in Fig. 4(c). Then we have

$$b_R = b_L \ . \qquad (6.10)$$

With (5.19), (5.20), and (6.10), the output mode a_L is expressed by (Yurke, 1984)

$$a_L = \sqrt{G'} a_R + e^{i\varphi} \sqrt{G'-1} a_R^{+} \ , \qquad (6.11)$$

where the overall signal gain $\sqrt{G'}$ is now given by

$$\sqrt{G'} = \frac{G}{2-G} \ . \qquad (6.12)$$

Note that (6.11) has the same form as (6.5) for the degenerate parametric amplifier, when $\varphi = 0$.

VII. NOISE OF PHOTON NUMBER DETECTORS

A. Photomultiplier

The normalized variance of the output photon number fluctuations for a photomultiplier is (Schockley and Pierce, 1938)

$$\langle \Delta n^2 \rangle_{\rm eq} = \langle \Delta n^2 \rangle + \langle n \rangle \delta^2 \frac{\left[1 - \dfrac{1}{m}\right]}{\langle g \rangle (\langle g \rangle - 1)} \ . \qquad (7.1)$$

Here $\langle \Delta n^2 \rangle$ is the variance of an input signal, $\langle g \rangle$ is the signal gain at each multiplication stage, $M = \langle g \rangle^m$ is the overall gain, m is the number of multiplication stages and $\delta^2=\langle g^2\rangle-\langle g\rangle^2$ is the mean-square deviation of gain. In the limit of high gain, M, $\langle g\rangle \gg 1$ and $\delta^2=\langle g\rangle$ (the secondary electron emission rate is Poissonian), the second term of (7.1) vanishes. Therefore, a photomultiplier can approach an ideal photon counter. Note that unity quantum efficiency is assumed in (7.1).

B. Avalanche photodiode

The normalized variance of the output photon number fluctuations for an avalanche photodiode is (McIntyre, 1965)

$$\langle \Delta n^2\rangle_{\rm eq}=\langle \Delta n^2\rangle+\langle n\rangle\left[k(M-1)+\left[1-\frac{1}{M}\right](1-k)\right]. \tag{7.2}$$

Here k is the ratio of electron and hole ionization coefficients and M is the overall gain. In the ideal limit of $k=0$ and $M\gg 1$, the second term of (7.2) reduces to $\langle n\rangle$. Therefore, an avalanche photodiode does not approach an ideal photon counter, even when the ratio of ionization coefficients k is zero, and the quantum efficiency is unity.

A new structure for a noise-free avalanche photodiode was proposed (Capasso, 1983) that has spatially confined avalanche regions and, therefore, has the same characteristics as a photomultiplier.

VIII. QUANTIZATION OF SHANNON'S CHANNEL CAPACITY

The maximum amount of information that can be carried by quantized electromagnetic waves is derived in this section. The derivation is based on the negentropy principle of information by Brillouin (1956) and the sampling theorem of Nyquist (1928). The result can be applied to the information capacity of any kind of band-limited bosonic mode.

A. The negentropy of information

The entropy S of a single mode of an electromagnetic wave with an average number of photons $\langle n\rangle$ is given by the maximum of the expression

$$\frac{S}{k}\equiv H=-\sum_n P(n)\ln P(n) \tag{8.1}$$

under the constraints

$$\sum_n P(n)=1 \tag{8.2a}$$

and

$$\sum_n nP(n)=\langle n\rangle . \tag{8.2b}$$

When the maximization is carried out one finds (Landau and Lifshitz, 1959) that $P(n)$ should satisfy the distribution

$$P(n)=\left[\frac{\langle n\rangle}{1+\langle n\rangle}\right]^n\frac{1}{1+\langle n\rangle} . \tag{8.3a}$$

The maximum entropy for this distribution is

$$H_{\rm max}=\langle n\rangle\ln\left[1+\frac{1}{\langle n\rangle}\right]+\ln(1+\langle n\rangle) . \tag{8.3b}$$

This is the maximum entropy of a system of $\langle n\rangle$ photons normalized by the Boltzmann constant.

The average number of photons is

$$\langle n\rangle=\langle n_s\rangle+\langle n_t\rangle , \tag{8.4}$$

where $\langle n_t\rangle$ is the number of thermal photons and is obtained from

$$\langle n_t\rangle=\frac{1}{\exp(\hbar\omega/kT)-1} \tag{8.5}$$

and if the average power of the signal is P, $\langle n_s\rangle$ is given by

$$\langle n_s\rangle=\frac{P\tau}{\hbar\omega} .$$

Here τ^{-1} is the arrival rate of the independent modes which constitute the signal wave. We shall define the arrival rate as a function of the channel bandwidth B in the next section.

According to the negentropy principle of information (Brillouin, 1956), the maximum amount of information I that can be extracted from signal states is equal to the difference between the total entropy (8.3) and the (residual) noise entropy which is

$$H=\ln(1+\langle n_t\rangle)+\langle n_t\rangle\ln\left[1+\frac{1}{\langle n_t\rangle}\right] . \tag{8.6}$$

Taking the difference between (8.5) and (8.6) one obtains

$$\begin{aligned} I\equiv H_{\rm max}-H=&(\langle n_s\rangle+\langle n_t\rangle)\ln\left[1+\frac{1}{\langle n_s\rangle+\langle n_t\rangle}\right]\\ &+\ln\left[1+\frac{\langle n_s\rangle}{1+\langle n_t\rangle}\right]\\ &-\langle n_t\rangle\ln\left[1+\frac{1}{\langle n_t\rangle}\right] . \end{aligned} \tag{8.7}$$

B. Quantum-mechanical channel capacity and information capacity of single photon

The channel capacity is the product of the maximum amount of information I of each mode and the arrival rate $B=1/\tau$, where τ is the time-interval occupied by each mode:

$$C=B\left[(\langle n_s\rangle+\langle n_t\rangle)\ln\left[1+\frac{1}{\langle n_s\rangle+\langle n_t\rangle}\right]\right.$$
$$\left.+\ln\left[1+\frac{\langle n_s\rangle}{1+\langle n_t\rangle}\right]-\langle n_t\rangle\ln\left[1+\frac{1}{\langle n_t\rangle}\right]\right]. \tag{8.8}$$

In the limiting case of low temperature, $\langle n_t\rangle \ll \langle n_s\rangle$, the channel capacity reduces to

$$C=B\left[\ln(1+\langle n_s\rangle)+\langle n_s\rangle\ln\left[1+\frac{1}{\langle n_s\rangle}\right]\right]. \tag{8.9}$$

This photon channel capacity was derived by Stern (1960), Gordon (1962), Levedev and Levitin (1963), Takahashi (1965), and She (1968). The first term is the product of the mode arrival rate B and the logarithm of the number of photons per mode plus one. This part of the information predominates in the limit of large photon number and is associated with the wave nature of the photons. The second term is the product of the arrival rate of photons, $B\langle n_s\rangle$, and the logarithm of the number of modes per photon $1/\langle n_s\rangle$ plus one. This information is carried by the particle property of the photons and predominates when the number of photons per mode is smaller than unity.

The information which can be carried by a single photon is given by

$$\frac{C}{B\langle n_s\rangle}=\ln\left[1+\frac{1}{\langle n_s\rangle}\right]+\frac{1}{\langle n_s\rangle}\ln(1+\langle n_s\rangle). \tag{8.10}$$

In the limit of small photon number, $\langle n_s\rangle \ll 1$, (8.10) goes to infinity. That is, a single photon can transmit an infinite amount of information in this limit.

In the opposite limiting case of $\langle n_s\rangle \ll \langle n_t\rangle$, the channel capacity reduces to

$$C=B\langle n_s\rangle\ln\left[1+\frac{1}{\langle n_t\rangle}\right]=B\langle n_s\rangle\frac{\hbar\omega}{kT}. \tag{8.11}$$

The information which can be transmitted by a single photon is now finite:

$$\frac{C}{B\langle n_s\rangle}=\frac{\hbar\omega}{kT}. \tag{8.12}$$

This number corresponds to about 32 nat/photon (46 bit/photon) for a wavelength of 1.5 μm and a temperature of 300 K. Pierce (1978) obtained (8.12) for the specific scheme using a coherent state, pulse position modulation (PPM), and photon counting detection. The derivation of (8.10) and (8.12) shown above does not assume any specific modulation-demodulation scheme and quantum state and, therefore, sets an upper limit on the information capacity by any selection of quantum states, modulation schemes and detection types.

The fact that the information per photon can go to infinity in the absence of thermal radiation may be made plausible as follows. If the wave packet of the photon can be sent at arbitrary times without other accompanying photons within a very long time interval (PPM), then this single photon can carry an amount of information which increases logarithmically with the ratio of the duration of the wave packet to the time interval. In order to determine the arrival of the wave packet, however, a bandwidth is required which is much greater than the arrival rate of the photons.

The ultimate information capacity (8.12) is realized only when the channel capacity per unit bandwidth (channel efficiency) is sacrificed enormously:

$$\frac{C}{B}=\langle n_s\rangle\frac{\hbar\omega}{kT}\ll\langle n_t\rangle\frac{\hbar\omega}{kT}\simeq 4\times10^{-13}. \tag{8.13}$$

A system with signal transmission via properly prepared photon number eigenstates, and detection of photons with ideal photon counters that are, according to Sec. VII free of noise, achieves the channel capacity (8.9) for $\langle n_t\rangle \ll \langle n_s\rangle$. We used photon number states for the purposes of analysis. If the transmission is accomplished by preparation of states other than photon number states, the channel capacity may, and in general does, differ from (8.9). Yet it can never exceed (8.9) which provides the ideal limit. We shall look at this issue in greater detail in the next sections.

C. Sampling theorem and bandwidth

Because the proper interpretation of the channel bandwidth B and the arrival rate τ^{-1} of an independent mode is crucial to the results of this paper, we review here briefly the considerations that lead to the relation $B=1/\tau$.

Suppose one starts with a baseband signal with a square spectrum centered around zero frequency. The Fourier transform of the flat spectrum extending from $\omega=-\pi B$ to $\omega=\pi B$ is proportional to $\sin(\pi Bt/\pi Bt)$ as shown in Fig. 7(a). A sequence of Nyquist functions displaced by

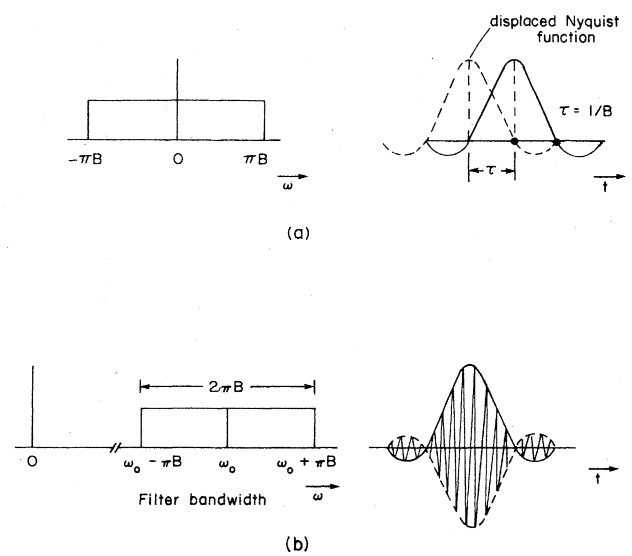

FIG. 7. The spectra of a baseband time function and a carrier time function and their Fourier transforms.

$\tau=1/B$ reproduces fully any bandwidth-limited function confined to the spectral width $B/2$ (positive part of spectrum) (Nyquist, 1928). Each Nyquist function corresponds to a mode which carries the information (8.7).

If the spectrum is centered at $\omega_0 \gg 2\pi B$, then the Nyquist function appears as in Fig. 7(b). A displacement by $\tau=1/B$ of the Nyquist function produces a new Nyquist function orthogonal to the original one. Again, one may represent a general band-limited time function as a superposition of Nyquist functions displaced by $\tau=1/B$.

Each mode has two degrees of freedom, phase, and amplitude, reflected in the fact that the amplitudes of the Nyquist function are complex. However, this does not imply that the channel capacity is given by multiplying (8.7) with the arrival rate times the number of degrees of freedom $2B$. The reason for this is the following: The derivation of (8.8) assumes that the states can be prepared for transmission and then measured in a noise-free manner. This is possible, in principle, with number states, the measurement of which can be performed ideally with no noise associated with the measurement as seen in Sec. VII. The measurement of phase and amplitude cannot be performed simultaneously without the introduction of additional uncertainty (noise) beyond that associated with the uncertainty principle as discussed in the theory of simultaneous measurements of two quadrature components. Therefore, Eq. (8.8) implies the use of one observable in the transmission and reception, one degree of freedom.

In the next section we present the channel capacity for the channel in which both amplitude and phase are used for transmission.

IX. CHANNEL CAPACITY OF CONTINUUM CHANNEL

When the transmission of information is accomplished by preparation of coherent states, or squeezed states, the transmitted observables can assume a continuum of values. Of course, both the prepared states, and the measurement of two complementary observables are subject to noise which imposes an upper limit on the achievable channel capacity.

A. Joint conditional probability and mutual information

Let us suppose that the two quadrature components a_1 and a_2 of the signal are used for the transmission of information. The joint conditional probability of the output events, given $\langle a_1\rangle$ and $\langle a_2\rangle$ and the measurement uncertainties $\langle\sigma_1^2\rangle$ and $\langle\sigma_2^2\rangle$, is

$$P(x,y;\langle a_1\rangle,\langle a_2\rangle)=\frac{1}{2\pi(\langle\sigma_1^2\rangle\langle\sigma_2^2\rangle)^{1/2}}\exp\left[-\frac{(x-\langle a_1\rangle)^2}{2\langle\sigma_1^2\rangle}-\frac{(y-\langle a_2\rangle)^2}{2\langle\sigma_2^2\rangle}\right]. \tag{9.1}$$

The noise (or residual) entropy H for the above signal with $\langle a_1\rangle$ and $\langle a_2\rangle$ is

$$\begin{aligned}H&\equiv-\int\int P(x,y;\langle a_1\rangle,\langle a_2\rangle)\ln[P(x,y;\langle a_1\rangle,\langle a_2\rangle)]dx\,dy\\&=\ln 2\pi+\tfrac{1}{2}\ln\langle\sigma_1^2\rangle+\tfrac{1}{2}\ln\langle\sigma_2^2\rangle+1\ .\end{aligned} \tag{9.2}$$

The maximum entropy that the signal can possess is calculated for a Gaussian distribution with zero mean and variances $\langle\Sigma_1^2\rangle$ and $\langle\Sigma_2^2\rangle$:

$$P(x,y)=\frac{1}{2\pi(\langle\Sigma_1^2\rangle\langle\Sigma_2^2\rangle)^{1/2}}\exp\left[-\frac{x^2}{2\langle\Sigma_1^2\rangle}-\frac{y^2}{2\langle\Sigma_2^2\rangle}\right]. \tag{9.3}$$

The maximum entropy is then given by

$$H_{\max}=\ln 2\pi+\tfrac{1}{2}\ln\langle\Sigma_1^2\rangle+\tfrac{1}{2}\ln\langle\Sigma_2^2\rangle+1\ . \tag{9.4}$$

When the signal and noise are independent, the channel capacity is

$$C=BI=B(H_{\max}-H)=B\left[\frac{1}{2}\ln\frac{\Sigma_1^2}{\langle\sigma_1^2\rangle}+\frac{1}{2}\ln\frac{\Sigma_2^2}{\langle\sigma_2^2\rangle}\right]. \tag{9.5}$$

The variances of the noise $\langle\sigma_1^2\rangle$ and $\langle\sigma_2^2\rangle$ are set by the quantum noises of the signal and detector.

B. Channel capacity of coherent state

Although a coherent state has the two quadrature fluctuations $\langle\Delta a_1^2\rangle=\langle\Delta a_2^2\rangle=\frac{1}{4}$, the simultaneous measurement of the two quadrature components introduces additional noise according to (4.11)

$$\langle\sigma_1^2\rangle=\langle\Delta a_{1,\mathrm{eff}}^2\rangle=\frac{1}{4}+\frac{\langle\Delta F_1^2\rangle}{G_1}\ , \tag{9.6}$$

$$\langle\sigma_2^2\rangle=\langle\Delta a_{2,\mathrm{eff}}^2\rangle=\frac{1}{4}+\frac{\langle\Delta F_2^2\rangle}{G_2}\ . \tag{9.7}$$

The minimum uncertainty product for such additional noise was shown to be

$$\frac{\langle\Delta F_1^2\rangle}{G_1}\frac{\langle\Delta F_2^2\rangle}{G_2}=\frac{1}{16}\ . \tag{9.8}$$

The minimization of the noise (or residual) entropy (9.2) under the constraints (9.6)–(9.8) requires $\langle\Delta F_1^2\rangle/G_1=\langle\Delta F_2^2\rangle/G_2=\frac{1}{4}$. The maximum entropy (9.4) is calculated by the relation

$$\langle \Sigma_1^2 \rangle + \langle \Sigma_2^2 \rangle = \langle n_s \rangle + 1 \ . \tag{9.9}$$

Here the additive 1 on the right-hand side of (9.9) stems from the quantum noise of the signal and the internal-mode of the detector. The maximization of the channel capacity (9.5) under these constraints is

$$C = B \ln(1 + \langle n_s \rangle) \ . \tag{9.10}$$

The signal modulation should satisfy the condition $\langle \Sigma_1^2 \rangle = \langle \Sigma_2^2 \rangle = \frac{1}{2}(\langle n_s \rangle + 1)$. The maximum number of bits which can be recovered by the combination of a coherent state and a simultaneous detector is 1.44 bit per photon.

C. Channel capacity for single quadrature measurement of squeezed state

For a squeezed state with $\langle \Delta a_1^2 \rangle \ll \langle \Delta a_2^2 \rangle$, the maximization of (9.5) requires that all the useful signal energy is used for modulating $\langle a_1 \rangle$. The signal consists of a sequence of squeezed states with different $\langle a_1 \rangle$ values but the same $\langle a_2 \rangle$ value of zero. Thermal noise is neglected as compared with the quantum noise $\langle \Delta a_1^2 \rangle$. If each Nyquist mode has the same noise distributions $\langle \Delta a_1^2 \rangle$ and $\langle \Delta a_2^2 \rangle$ but has different $\langle a_1 \rangle$ values, the ensemble of these Nyquist modes has the Gaussian distribution:

$$P(x,y) = \frac{1}{2\pi(\langle \Sigma_1^2 \rangle \langle \Sigma_2^2 \rangle)^{1/2}} \exp\left[-\frac{x^2}{2\langle \Sigma_1^2 \rangle} - \frac{y^2}{2\langle \Sigma_2^2 \rangle} \right] , \tag{9.11}$$

$$\langle \Sigma_1^2 \rangle = \langle n_s \rangle + \tfrac{1}{2} - \langle \Delta a_2^2 \rangle + \langle F_1^2 \rangle \ , \tag{9.12}$$

$$\langle \Sigma_2^2 \rangle = \langle \Delta a_2^2 \rangle + \langle F_2^2 \rangle \ , \tag{9.13}$$

where $\langle F_1^2 \rangle$ and $\langle F_2^2 \rangle$ is the quantum noise imposed by the detector, which satisfies $\langle F_1^2 \rangle \langle F_2^2 \rangle = \frac{1}{16}$. The maximum entropy is calculated for this Gaussian distribution. After the measurement, the signal state is left in the state which satisfies

$$P(x,y;\langle a_1 \rangle, \langle a_1 \rangle = 0) = \frac{1}{2\pi(\langle \sigma_1^2 \rangle \langle \sigma_2^2 \rangle)^{1/2}} \exp\left[-\frac{(x - \langle a_1 \rangle)^2}{2\langle \sigma_1^2 \rangle} - \frac{y^2}{2\langle \sigma_2^2 \rangle} \right] , \tag{9.14}$$

$$\langle \sigma_1^2 \rangle = \langle \Delta a_1^2 \rangle + \langle F_1^2 \rangle \ , \tag{9.15}$$

$$\langle \sigma_2^2 \rangle = \langle \Delta a_2^2 \rangle + \langle F_2^2 \rangle \ . \tag{9.16}$$

The residual entropy is calculated for this distribution. The information is

$$I \equiv H_{\max} - H = \tfrac{1}{2}\ln\langle \Sigma_1^2 \rangle + \tfrac{1}{2}\ln\langle \Sigma_2^2 \rangle - \tfrac{1}{2}\ln\langle \sigma_1^2 \rangle - \tfrac{1}{2}\ln\langle \sigma_2^2 \rangle = \tfrac{1}{2}\ln\frac{\langle n_s \rangle + \frac{1}{2} - \langle \Delta a_2^2 \rangle + \langle F_1^2 \rangle}{\langle \Delta a_1^2 \rangle + \langle F_1^2 \rangle} \ . \tag{9.17}$$

This value becomes a maximum when

$$\langle \Delta a_1^2 \rangle = \frac{1}{4(2\langle n_s \rangle + 1)} \ , \tag{9.18}$$

$$\langle \Delta a_2^2 \rangle = \frac{2\langle n_s \rangle + 1}{4} \ , \tag{9.19}$$

$$\langle F_1^2 \rangle = 0 \quad \text{(single quadrature measurement)} \ . \tag{9.20}$$

Equations (9.18) and (9.19) show that the overall state which maximizes the signal-to-noise ratio (Yuen and Shapiro, 1980) is optimum in terms of the channel capacity.

The maximum channel capacity is

$$C = B \ln(1 + 2\langle n_s \rangle) \ . \tag{9.21}$$

This result was obtained by Yuen (1983). The maximum number of bits which can be recovered by the combination of a squeezed state and a single quadrature measurement detector is 2.88 bit per photon.

The channel capacities C, normalized by the channel bandwidth B, are called channel efficiencies and are plotted in Fig. 8 as functions of the average photon number $\langle n_s \rangle$ per mode. The channel efficiencies for the simultaneous measurement of coherent states and for the single quadrature measurement of squeezed states are lower than those for a photon counting measurement of photon number states. The drop off in the very small photon number, $\langle n_s \rangle \ll 1$, is not inherent in the signal quantum states but stems from the inefficiency of the detectors.

D. Channel capacity for photon counting measurement of coherent state

Let us consider finally the case of a signal in a coherent state, detected by a photon counter. The photoelectron statistics for the given average photon number $s = \langle n \rangle$

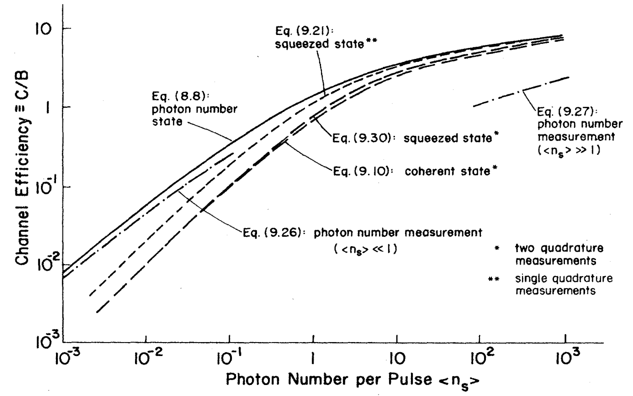

FIG. 8. The quantum-mechanical form of Shannon's channel capacity (8.8), the channel capacities recovered by the two quadrature measurements of a coherent state (9.10), of a squeezed state (9.30), by a single quadrature measurement of a squeezed state (9.21), and by photon counting of a coherent state for a small photon number (9.26), and for a large photon number (9.27).

are Poissonian:

$$P(n)=\frac{e^{-s}s^n}{n!} . \tag{9.22}$$

When the average number of signal photons per mode is much smaller than unity, the information per mode is given by (Gordon, 1962)

$$I=H(y)-H_x(y) , \tag{9.23}$$

where $H(y)$ is the total information and is given by the probability $P(0)$ that no photons are received and $P(1)$, the probability that one or more photons are received:

$$\begin{aligned} H(y)&=-P(0)\ln P(0)-P(1)\ln P(1) \\ &=-Q(1-e^{-s})\ln[Q(1-e^{-s})] \\ &\quad -[1-Q(1-e^{-s})]\ln[1-Q(1-e^{-s})] , \end{aligned} \tag{9.24}$$

where Q is the probability of sending photons (on state) and $1-Q$ is the probability of sending no photons (off state). The conditional probabilities are

$$P(1)_{\rm on}=1-e^{-s} , \quad P(0)_{\rm on}=e^{-s} ,$$

$$P(1)_{\rm off}=0 , \quad P(0)_{\rm off}=1 .$$

$H_x(y)$ is the conditional entropy per mode when the transmitter symbol (on or off) is known

$$H_x(y)=-Q[e^{-s}\ln e^{-s}+(1-e^{-s})\ln(1-e^{-s})] . \tag{9.25}$$

The maximization of (9.23)–(9.25) gives the following channel capacity in the limit of $\langle n_s \rangle = Qs \ll 1$,

$$C=BsQ\ln\left[\frac{1}{sQ}\right]=B\langle n_s\rangle\ln\left[\frac{1}{\langle n_s\rangle}\right] . \tag{9.26}$$

The channel capacity approaches the upper bound of the photon channel capacity (8.8), as the average number of signal photons goes to zero as shown in Fig. 7.

When the average number of signal photons $\langle n_s \rangle$ per mode is much larger than unity, the channel capacity is (Gordon, 1962)

$$C \simeq \frac{B}{2}\ln\langle n_s\rangle . \tag{9.27}$$

In the limit of $\langle n_s \rangle \gg 1$, the channel capacity recovered by a photon counter is one-half of the photon channel capacity (8.8). This can be understood as follows. When the average number of signal photons $\langle n_s \rangle$ is much larger than unity, half of the information is carried in the form of photon number and the remaining half in the form of phase, which is rejected in the photon counter. When $\langle n_s \rangle$ is much smaller than unity, on the other hand, the photon phase cannot be defined (Carruthers and Neito, 1968). In this range, all the information may be carried in the form of photon number which can be recovered by an ideal photon counter.

Although an amplitude-squeezed state has sub-Poissonian photon statistics, the deviation from the Poisson distribution is not large (Yuen, 1976). The channel capacity for a photon counting measurement of (amplitude) squeezed states is bounded by (8.8), and (9.26), and (9.27).

E. Channel capacity with beam splitter followed by single quantum measurement detectors

The signal and noise quantities for a beam splitter with $\varepsilon=\frac{1}{2}$ are given in (5.28) to (5.31). The squares of the mean amplitudes are

$$\langle a_{1,\rm eff}\rangle^2=\langle a_{2,\rm eff}\rangle^2=\frac{\langle n_s\rangle}{2} \tag{9.28}$$

and fluctuations

$$\langle \Delta a_{i1}^2\rangle=\langle \Delta a_{i2}^2\rangle=\langle \Delta a_{s1}^2\rangle=\langle \Delta a_{s2}^2\rangle=\tfrac{1}{4} . \tag{9.29}$$

When these values are introduced in (9.5), one obtains the channel capacity (9.10) with $\varepsilon=\frac{1}{2}$. The same result is achieved by heterodyne detection with a coherent state. The excess noise required for a simultaneous measurement of two noncommuting variables came from the "open port" of the beam splitter.

When squeezed states are used, one has the choice of putting all the information into the in-phase signal channel and decouple from the open port, setting $\varepsilon=1$. In this limit one obtains the result (9.21) which is also obtained for the heterodyne detection of a squeezed state. Of course, this limit is quite uninteresting, because it is a trivial example of the channel capacity with the signal carried by only one of two noncommuting variables and detected by an ideal noise free detector.

A case of greater interest is when one puts all the information into the in-phase signal channel, and uses a beam splitter with $\varepsilon=\frac{1}{2}$. In the signal channel one uses squeezed states given by (9.18) and (9.19), and the open port is excited by an equally squeezed ground state. Then the channel capacity becomes

$$C=\frac{B}{2}\ln[1+2\langle n_s\rangle(1+\langle n_s\rangle)] . \tag{9.30}$$

The same result is obtained for heterodyne detection with analogously prepared states in the signal channel and the idler channel, respectively.

X. DISCUSSION

This paper reviewed the signal and noise properties of various detection schemes. The fact that a simultaneous measurement of two noncommuting observables leads to additional noise was emphasized and was demonstrated as a matter of principle (two simultaneously measurable quantities must be made to commute and thus require additive noise) and with the aid of particular device equations. The quantum noise associated with the signal preparation and the measurement process sets an upper bound for the channel capacity. This upper bound was derived generally from the negentropy principle. Any signal-state preparation and measurement scheme can only lead to a channel capacity that is lower, or at best equal, to the limiting channel capacity.

The internal-mode fluctuations, the channel capacity, and the information capacity for various detectors are summarized in Table I. All detectors capable of a simultaneous measurement of two quadrature components have inevitable internal-mode fluctuations, which increase the uncertainty product of the measurement by 3 dB from the Heisenberg uncertainty principle. Such detectors can recover the highest channel capacity from a coherent state, if the average number of photons is much larger than unity. The ultimate information capacity is 1.44 bit per photon. All ideal detectors for a single quadrature measurement are free from internal-mode fluctuations, and realize the Heisenberg uncertainty principle. Such detectors can recover the highest channel capacity from a squeezed state, again if the average number of photons is much larger than unity. The ultimate information capacity is 2.88 bit per photon. A photon counter, on the other hand, can reach the photon channel capacity if a photon

TABLE I. Classification of various optical measuring schemes, internal noise, channel capacity C, and information capacity $C/B\langle n_s\rangle$ of a single photon. $|0\rangle$, zero-point fluctuation; $|0\rangle_\beta$, squeezed zero-point fluctuation. CS, coherent state; SS, squeezed state; PNS, photon number state.

Class	Detectors	Internal-mode (fluctuation)	Channel capacity C information capacity $C/B\langle n_s\rangle$
Balanced detector pair	Heterodyning	$\lvert 0\rangle$ at image band	
	beam splitter	$\lvert 0\rangle$ from open port	
Phase insensitive amplifier	Parametric amplifier	$\lvert 0\rangle$ at idler band	$C=B\ln(1+\langle n_s\rangle)$
	Degenerate 4-wave mixer	$\lvert 0\rangle$ from open port	
	Raman amplifier	Optical phonon	$\frac{C}{B\langle n_s\rangle}\times 1.44$ bit
	Brillouin amplifier	Acoustic phonon	
	Laser amplifier	Dipole moment	
QND meas.	Optical Kerr medium	Probe wave	(CS)
Squeezed internal mode	Heterodyning	$\lvert 0\rangle_\beta$ at image band	$C=\frac{B}{2}\ln[1+2\langle n_s\rangle(\langle n_s\rangle+1)]$
	Beam splitter	$\lvert 0\rangle_\beta$ from open port	
	Parametric amplifier	$\lvert 0\rangle_\beta$ at idler band	$\frac{C}{B\langle n_s\rangle}\times 1.44$ bit
	Degenerate 4-wave mixer	$\lvert 0\rangle_\beta$ from open port	
	Optical Kerr medium	Probe wave (nonsqueezed)	(SS)
Single quadrature measurement	Degenerate heterodyning		$C=\frac{B}{2}\ln(1+4\langle n_s\rangle)$
	Degenerate parametric amplifier		$\frac{C}{B\langle n_s\rangle}\rightarrow 2.88$ bit (CS)
	Cavity degenerate 4-wave mixer		$C=B\ln(1+2n_s)$ $\frac{C}{B\langle n_s\rangle}\rightarrow 2.88$ bit (SS)
Photon counting measurement	Photomultiplier		$C=B\left[\ln(1+\langle n_s\rangle)+\langle n_s\rangle\ln\left[1+\frac{1}{\langle n_s\rangle}\right]\right]$ $\frac{C}{B\langle n_s\rangle}\rightarrow\infty$ (PNS)
	Avalanche photodiode		$C=B\langle n_s\rangle\ln\left[\frac{1}{\langle n_s\rangle}\right]$ $\frac{C}{B\langle n_s\rangle}\rightarrow\infty$ (CS)

number state is prepared, and also recovers the highest channel capacity from a coherent state, if the average number of photons is much smaller than unity. The efficiency is, however, 50% when the average number is much larger than unity.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation under Grant No. 8305448-AO1-ECS. One of the authors (Y.Y.) wants to thank F. Kanaya for his useful discussion.

REFERENCES

Aharonov, Y., and D. Bohm, 1961, Phys. Rev. **122**, 1649.

Aharonov, Y., and D. Bohm, 1964, Phys. Rev. **134**, B1417.

Arthurs, E., and J. L. Kelly, Jr., 1965, Bell Syst. Tech. J. **44**, 725.

Bohm, D., 1951, *Quantum Theory* (Prentice-Hall, Englewood Cliffs).

Bohr, N., 1935, Phys. Rev. **48**, 696.

Bondurant, R. S., Kumar, J. H. Shapiro, and M. Maeda, 1984, Phys. Rev. A **30**, 343.

Bondurant, R. S., and J. H. Shapiro, 1984, Phys. Rev. D **30**, 2548.

Braginsky, V. B., C. M. Caves, and K. S. Thorne, 1977, Phys. Rev. D **15**, 2047.

Braginsky, V. B., and B. Y. Khalili, 1983, Sov. Phys.—JETP **57**, 1124.

Braginsky, V. B., Y. I. Vorontsov, and K. S. Thorne, 1980, Science **209**, 547.

Braginsky, V. B., and S. P. Vyatchamin, 1981, Dokl, Akad. Nauk SSSR **257**, 570 [Sov. Phys. Dokl. **26**, 686 (1981)].

Braginsky, V. B., and S. P. Vyatchamin, 1982, Dokl, Akad. Nauk SSSR **264**, 1136 [Sov. Phys. Dokl. **27**, 478 (1982)].

Brillouin, L., 1956, *Science and Information Theory* (Academic, New York).

Capasso, F., 1983, IEEE Trans. Electron. Devices **ED-30**, 381.

Carruthers, P., and M. M. Neito, 1968, Rev. Mod. Phys. **40**, 411.

Caves, C. M., 1982, Phys. Rev. D **26**, 1817.

Caves, C. M., K. S. Thorne, R. W. P. Drever, V. D. Sandberg, and M. Zimmermann, 1980, Rev. Mod. Phys. **52**, 341.

Chan, V. W. S., 1981, Proc. Soc. Photo-Opt. Instrum. Eng. **295**, 10.

Dorschner, T. A., H. A. Haus, M. Holz, I. W. Smith, and H. Statz, 1980, IEEE J. Quantum Electron. **QE-16**, 1376.

Einstein, A., B. Podolsky, and N. Rosen, 1935, Phys. Rev. **47**, 777.

Fano, R. M., 1961, *Transmission of Information* (MIT, Cambridge, Mass.).

Glauber, R. J., 1951, Phys. Rev. **84**, 395.

Glauber, R. J., 1963, Phys. Rev. **131**, 2766.

Gordon, J. P., 1962, Proc. IRE **50**, 1898.

Haken, H., *Light and Matter Ic* (Vol. XXV/2c of *Encyclopedia of Physics*), edited by S. Flügge (Springer, Berlin-Heidelberg-New York).

Haus, H. A., and J. A. Mullen, 1962, Phys. Rev. **128**, 2407.

Haus, H. A., and C. H. Townes, 1962, Proc. IRE **50**, 1544.

Haus, H. A., and Y. Yamamoto, 1984, Phys. Rev. A **29**, 1261.

Haus, H. A., and Y. Yamamoto, 1986, Phys. Rev. A, in press.

Helstrom, C. W., 1976, *Quantum Detection and Estimation Theory* (Academic, New York).

Imoto, N., H. A. Haus, and Y. Yamamoto, 1985, Phys. Rev. A **32**, 2287.

Kubo, R., 1965, *Statistical Mechanics; an Advanced Course with Problems and Solutions* (North-Holland, Amsterdam).

Landau, L. D., and E. M. Lifshitz, 1959, *Statistical Physics* (Pergamon, London/Paris), Vol. 5.

Levedev, D. S., and L. B. Levitin, 1963, Dokl. Akad. Nauk SSSR **149**, 1299 [Sov. Phys. Dokl. **8**, 377 (1963)].

Levenson, M., R. M. Shelby, A. Aspect, M. Reid, and D. F. Walls, 1985, Phys. Rev. A **32**, 1550.

Louisell, W. H., 1973, *Quantum Statistical Properties of Radiation* (Wiley, New York).

Louisell, W. H., A. Yariv, and A. E. Siegman, 1961, Phys. Rev. **124**, 1646.

Machida, S., and Y. Yamamoto, 1986, Opt. Commun., in press.

McIntyre, R. J., 1965, IEEE Trans. Electron. Devices **ED-13**, 164.

Nyquist, H., 1928, AIEE Trans. **47**, 617.

Oliver, B. M., 1962, Proc. IRE **50**, 1545.

Personick, S. D., 1971, Bell Syst. Tech. J. **50**, 213.

Pierce, J. R., 1978, IEEE Trans. Commun. **COM-26**, 1819.

Pierce, J. R., E. D. Posner, and E. R. Rodemich, 1981, IEEE Trans. Inf. Theory **IT-27**, 61.

Sanders, G. A., M. G. Prentiss, and S. Ezekiel, 1980, Opt. Lett. **6**, 569.

Sargent, M., III, M. O. Scully, and W. E. Lamb, Jr., 1974, *Laser Physics* (Addison-Wesley, Reading, Mass.).

Schawlow, A., and C. H. Townes, 1958, Phys. Rev. **112**, 1940.

Shannon, C. E., 1948, Bell Syst. Tech. J. **27**, 379.

Shapiro, J. H., 1985, IEEE J. Quantum Electron **QE-21**, 237.

She, C. Y., 1968, IEEE Trans. Inf. Theory **IT-14**, 32.

She, C. Y., and H. Heffner, 1966, Phys. Rev. **152**, 1103.

Shimoda, K., H. Takahashi, and C. H. Townes, 1957, J. Phys. Soc. Jpn. **12**, 686.

Shockley, W., and J. Pierce, 1938, Proc. IRE **26**, 321.

Short, R., and L. Mandel, 1983, Phys. Rev. Lett. **51**, 384.

Slusher, R. E., L. W. Hollberg, B. Yurke, J. C. Mertz, and J. F. Valley, 1985, Phys. Rev. Lett. **55**, 2409.

Stern, T. E., 1960, IEEE Trans. Inf. Theory **IT-6**, 435.

Stoler, D., 1971, Phys. Rev. D **4**, 1925.

Strandberg, H., 1963, Proc. IEEE **51**, 943.

Sudarshan, E. C. G., 1963, Phys. Rev. Lett. **10**, 277.

Takahashi, H., 1965, in *Advances in Communication Systems*, edited by A. V. Balakrishnan (Academic, New York), p. 277.

Thorne, K. S., 1980a, Rev. Mod. Phys. **52**, 285.

Thorne, K. S., 1980b, Rev. Mod. Phys. **52**, 299.

Thorne, K. S., R. W. P. Drever, C. M. Caves, M. Zimmermann, and V. D. Standberg, 1978, Phys. Rev. Lett. **40**, 667.

Unruh, W. G., 1978, Phys. Rev. D **18**, 1764.

Walls, D. F., 1983, Nature (London) **301**, 141.

Yamamoto, Y., 1980, IEEE J. Quantum Electron. **QE-16**, 1251.

Yamamoto, Y., N. Imot, and S. Machida, 1986, Phys. Rev. A., in press.

Yamamoto, Y., O. Nilsson, and S. Saito, 1984, Proceedings of the 3rd US-Japan Seminar on Quantum Electron. (Nara), unpublished.

Yariv, A., and W. H. Louisell, 1966, IEEE J. Quantum Electron. **QE-2**, 418.

Yuen, H. P., 1975, Proceedings of the 1975 Johns Hopkins Conference on Information Science and Systems.

Yuen, H. P., 1976, Phys. Rev. A **13**, 2226.
Yuen, H. P., 1983, Phys. Rev. Lett. **51**, 719.
Yuen, H. P., and J. H. Shapiro, 1979, Opt. Lett. **4**, 334.
Yuen, H. P., and J. H. Shapiro, 1980, IEEE Trans. Inf. Theory **IT-26**, 78.
Yurke, B., 1984, Phys. Rev. A **29**, 408.
Yurke, B., and J. S. Denker, 1984, Phys. Rev. A **29**, 1419.