Novel Quantum States and Measurements

Thesis by

Samuel L. Braunstein

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1988

(Submitted May 17, 1988)

Acknowledgments

Carl Caves deserves thanks for his supervision and friendship, especially for his support of the many, sometimes crazy, projects I have chosen to undertake.

Time out for me has often meant carrying on open-ended discussions, sipping coffee, eating German chocolate cake, · · · at the Red Door Café with my friend Mike Morris.

I thank my family, my parents in particular, for their personal support over the years, even with our long separation.

I have been funded during my years at Caltech by Research and Teaching Assistantships, the Caltech President's Fund, the National Science Foundation, and the Office of Naval Research. I gratefully acknowledge their support.

Abstract

i) The quantum mechanics of higher order parametric amplifiers is studied. It is shown that these devices produce meaningful quantum states; numerical calculations are performed that demonstrate the convergence of matrix elements associated with these states. Further, the correspondence between the classical and quantum evolution for these devices is studied and the differences explained by a kind of "quantum diffusion." Finally, the possibility of producing ordinary squeezed light with these devices is noted.

ii) The generation of squeezed light always involves some scheme that amounts to pumping electromagnetic modes at near twice their natural frequency. When the pump is itself treated quantum mechanically, extra noise is introduced that ultimately limits the amount of squeezing achievable. Detailed calculations are carried out in this regard for the parametric amplifier. It is found that the pump's initial phase noise is responsible for this limit.

iii) Quantum-mechanical measurements are usually described by applying the standard quantum rules to a measurement model. They can also be described by a formalism that uses mathematical objects called Effects and Operations. These two descriptions should be equivalent. D'Espagnat has raised a question about the usage of this formalism of Effects and Operations for repeated measurements. This question is cleared up, and the source of the discrepancy is given a simple interpretation.

iv) Usually, an inequality that is chained becomes a weaker inequality. Chaining the Bell inequality, however, leads to stronger violations by quantum mechanics. Further, a new kind of Bell inequality, based on the information obtained in a measurement, is derived. This information Bell inequality can be used to formulate tests of local realism in very general circumstances, *e.g.*, higher spin versions of the EPR experiment. These new inequalities yield an interpretation for the size of their violation and lead to the formulation of a hierarchy of Bell inequalities for which two-particle Bell inequalities play a special role.

Table of Contents

CHAPTER 1

Introduction

This thesis is a theoretical investigation whose general theme is the foundations of quantum theory – a topic that should and does concern every physicist. The work in this thesis consists of four research papers (all with co-authors) and neatly splits into two parts, which follow Chapter I: Part one deals with quantum noise, which is related to the quantum theory of measurement; and part two, truly at the foundations, deals with general measurements and the interpretation of quantum mechanics.

I chose this ordering for the presentation, *i.e.*, more practical issues followed by questions at the foundations of the theory, because I consider it the more logical (perhaps because it mirrors the development of my own research interests). After all, how can you understand the foundations without first understanding some of the implications of the theory? The remainder of this chapter will consist of two separate introductions, one for each of the two parts of this thesis. These introductions will motivate the works as well as outline the results obtained.

I.1 INTRODUCTION TO PART I

The research in part I of this thesis deals with squeezed light or generalizations to it. So I shall start by giving a brief tutorial on squeezed states of light. I will describe what a squeezed state is, how it is seen, made, and used. Following this tutorial, I introduce my own work on squeezing.

All light fields fluctuate, their amplitudes and phases being subject to various kinds of noise. Much of this noise in practical sources is due to environmental influences: The random rate of atomic excitation in a light source leads to amplitude fluctuations; random motion of the atoms can lead to frequency spreading, and collisions during the emission of quanta can cause fluctuations in the phase of the emitted light. Even if all these factors are eliminated by a suitably designed light source, we

will still be left with a more fundamental source of noise: noise that is due to quantum mechanics and the Heisenberg uncertainty principle. Although this noise is fundamental, that doesn't mean that we are prevented from making improvements. One practical way of reducing some aspects of this quantum noise is to "squeeze" the light.

I.2 WHAT IS SQUEEZED LIGHT?

In this section I shall illustrate the main ideas behind the theory of squeezed light. To simplify my discussion, I will henceforth be dealing only with the modes of the electromagnetic field that make up a nearly plane wave with a single polarization (*i.e.*, we will be dealing with the continuum of frequencies that contribute to an electromagnetic wave with a particular transverse mode structure). The electric field will be decomposed into two "quadrature" components that vary in time as $\cos(\omega_0 t)$ and $\sin(\omega_0 t)$. The electric field is then described by a mixture of these two components. The quantum-mechanical operators associated with these components, called quadrature phases, can be shown to be noncommuting Hermitian operators and hence will satisfy a Heisenberg uncertainty relation. When these quadrature phases have equal and minimum uncertainties ("noise"), they form a "coherent" state. The fact that the noises of the two components are equal implies that the electric field itself will have time stationary noise. One example of a coherent state is the vacuum state of the electric field; even the vacuum electric field has fluctuations! When the noises of the quadrature phases are unequal – yet their product is still a minimum – we obtain an ideal "squeezed" state of light. One quadrature of the electric field is then found to be less noisy than the vacuum, and the other, more noisy.

The electric field of the wave we are considering, at a fixed point in space, can be decomposed[1] as

$$E(t)=E_1(t)\cos(\omega_0 t)+E_2(t)\sin(\omega_0 t)\,, \tag{I.2.1}$$

where ω_0 is an optical frequency. The quantities E_1 and E_2, the quadrature phases relative to ω_0, allow a ''phase space'' description of the electric field. They are analogous to the position and momentum of a harmonic oscillator. If the electric field in Eq. (I.2.1) has a small enough bandwidth, and if we choose ω_0 to be an optical frequency at the center of this bandwidth, then our decomposition into $E_1(t)$ and $E_2(t)$ has removed the fast optical motion of the field, leaving only slow variations in the quadrature phases.

As quantum-mechanical operators, these quadrature phases satisfy the equal-time commutation relation

$$[E_1(t),E_2(t)]=i\frac{b\,\omega_0}{cA_q}\int_B d\omega\,\hbar\,, \tag{I.2.2}$$

where B is the bandwidth of the field, A_q serves as the cross-sectional area of the beam, and b is a units-dependent constant (*e.g.*, $b=4\pi$ in cgs Gaussian units). Because of this commutation relation, the quadrature phases have to satisfy the quantum-mechanical uncertainty relation

$$\Delta E_1\Delta E_2\geq\frac{b\,\omega_0}{cA_q}\int_B d\omega\frac{\hbar}{2}\,. \tag{I.2.3}$$

The light that produces interference patterns with the best contrast[2] is called coherent light. Coherent light is a minimum uncertainty state with the uncertainty of the two quadrature phases being equal and uncorrelated,

$$(\Delta E)_{\rm coh} \equiv \Delta E_1 = \Delta E_2 = \left[\frac{b\,\omega_0}{cA_q} \int_B d\,\omega \frac{\hbar}{2} \right]^{1/2} \tag{I.2.4}$$

In Fig. 1 I plot phase-space diagrams for several states of the electromagnetic field. I have given only E_1 a nonzero average field. Fluctuations along the E_1-direction will change the amplitude of the field (amplitude fluctuations), and fluctuations perpendicular to this direction will vary the phase (phase fluctuations). The quadratures E_1 and E_2 and their uncertainty, for a coherent state, are plotted in Fig. 1.(a). In this figure, the circle (the ''error circle'') represents the size of the fluctuations for each quadrature, and the position of the center of the circle represents their average value. Fig. 1.(b) shows the electric field in the time domain for the coherent state of Fig. 1.(a). The central line traces the average value for the electric field, and the upper and lower lines show the size of the fluctuations. The variance in the electric field can be worked out from Eq. (I.2.1) to be

$$(\Delta E)^2 = (\Delta E_1)^2 \cos^2(\omega_0 t) + (\Delta E_2)^2 \sin^2(\omega_0 t) + \Delta(E_1 E_2)\cos(\omega_0 t)\sin(\omega_0 t)\,,$$

$$= \frac{b\,\omega_0}{cA_q} \int_B d\,\omega \frac{\hbar}{2} = (\Delta E)^2_{\rm coh}\,, \tag{I.2.5}$$

for the coherent state, and is independent of time. Note that the correlation between the different quadratures,

$$\Delta(E_1 E_2) \equiv \langle E_1 E_2 + E_2 E_1 \rangle - 2\langle E_1 \rangle \langle E_2 \rangle\,, \tag{I.2.6}$$

is zero for a coherent state.

Squeezed light is also a minimum uncertainty state, but the uncertainties are not equal. I shall illustrate two choices of such a state where the quadratures are uncorrelated (all other possibilities correspond only to a ''rotation'' of the noise in the phase-

space picture). Fig. 1.(c) shows squeezed light with zero correlation between the quadratures and

$$\Delta E_1 = G(\Delta E)_{\text{coh}}$$

$$\Delta E_2 = \frac{(\Delta E)_{\text{coh}}}{G}, \qquad G > 1 . \tag{I.2.7}$$

The ellipse in this figure (the ''error ellipse'') indicates the size of fluctuations in the quadratures. With this choice of squeezing, and $<E_1> \neq 0$, we get reduced phase noise. Fig. 1.(d) shows the electric field for this configuration in the time domain; the variance of the electric field is

$$(\Delta E)^2 = \left[G^2 \cos^2(\omega_0 t) + \frac{\sin^2(\omega_0 t)}{G^2} \right] (\Delta E)^2_{\text{coh}} . \tag{I.2.8}$$

We see that the amplitude fluctuations (occurring at the crests of the electric field) are larger than those for the coherent state, while the phase (determined by the crossings of the electric field through zero) is more precisely defined. Such light is called phase-squeezed.

In the opposite configuration,

$$\Delta E_1 = \frac{(\Delta E)_{\text{coh}}}{G}$$

$$\Delta E_2 = G(\Delta E)_{\text{coh}}, \qquad G > 1 , \tag{I.2.9}$$

and

$$(\Delta E)^2 = \left[\frac{\cos^2(\omega_0 t)}{G^2} + G^2 \sin^2(\omega_0 t) \right] (\Delta E)^2_{\text{coh}} , \tag{I.2.10}$$

as shown in Figs. 1.(e) and (f), with $<E_1> \neq 0$ the phase fluctuations are larger and the amplitude fluctuations are smaller than for a coherent state. This arrangement of fluctuations relative to the average field is called amplitude-squeezed.

I.3 HOW DO WE SEE SQUEEZING?

How can we make use of this modification of the quantum noise? Well, if we had a phase-sensitive detector, then we could pick out that part of the electric field which has reduced fluctuations. One such detector is the balanced homodyne detector.[3] I shall give a brief description of its operation and will tell how the experimental signature for squeezing is recognized.

Fig. 2. shows a balanced homodyne detector. Its operation is simple enough for a classical description to suffice at the moment. A signal field E_S and a local oscillator (acting as a phase reference) field E_{LO} are combined at a 50–50 beam splitter. The sum and difference of these field amplitudes impinge on two separate photodetectors. Photodetectors produce a ''photocurrent'' that is proportional to the square of the amplitude of the field incident on them. Thus, the photodetectors shown in Fig. 2 will produce a pair of photocurrents I_1 and I_2:

$$I_1 \propto \frac{1}{2}(E_S - E_{LO})^2 , \tag{I.3.1}$$

$$I_2 \propto \frac{1}{2}(E_S + E_{LO})^2 . \tag{I.3.2}$$

Next, the difference of these currents is taken:

$$I_D \propto I_2 - I_1 \propto 2E_S E_{LO} . \tag{I.3.3}$$

The fastest observable variations in this differenced current will be for less than an

optical frequency. Suppose we decompose the signal field into its quadratures,

$$E_S(t) = E_{S1}(t)\cos(\omega_0 t) + E_{S2}(t)\sin(\omega_0 t)\,, \tag{I.3.4}$$

and choose the local oscillator's phase so that

$$E_{LO} = \mathcal{E}_{LO}\cos(\omega_0 t - \theta)\,. \tag{I.3.5}$$

Then the difference current averaged over times comparable with several optical cycles will be

$$I_D(t) \propto \mathcal{E}_{LO}\,[E_{S1}(t)\cos(\theta) + E_{S2}(t)\sin(\theta)]\,. \tag{I.3.6}$$

By sweeping the phase θ of the local oscillator, we can pick out either of the quadratures E_{S1} or E_{S2} or any linear combination of them.

In practice, the bandwidth available to I_D will be determined by the electronics and may be as large as many gigahertz. In this case, we will wish to look at smaller sections of the spectrum by putting the differenced current through a spectrum analyzer to get the power spectrum of the noise in I_D as a function of RF frequency.

A coherent state (*e.g.*, the vacuum) coming in as a signal has equal fluctuations in the two quadratures [see Eq. (I.2.3)], so its power spectrum will be independent of the local oscillator's phase; it will appear as a flat trace if we scan this phase. However, when a squeezed state is the input signal, the power spectrum will depend on the local oscillator's phase. Depending on which quadrature is picked out [via Eq. (I.3.6)], the power spectrum (at some fixed RF frequency) will rise above the vacuum noise level or drop below it.

Wu et al[4] have seen the biggest reduction below the vacuum level in an optical homodyne detector. They found a reduction of 63% below vacuum level, at a RF

frequency of 1.8 MHz over a bandwidth of 100 kHz. We show the plot of the rms noise voltage versus local oscillator phase from their paper in Fig. 3, *i.e.*, it is a plot of the square root of the noise power versus local oscillator phase at the fixed RF frequency of 1.8 MHz. The vacuum level was obtained by blocking the signal port to the homodyne detector; this level (which was found to be phase insensitive) is drawn as a dashed line. This plot illustrates the two important features of squeezed light: Its noise is phase sensitive, and it has a noise level that drops below vacuum.

I.4 HOW IS LIGHT SQUEEZED?

We have already seen that one main feature of squeezed light is its phase sensitive noise. In this section, I want to discuss how we can produce such light. Typically, sources of light we might find produce noise that is phase-insensitive. For example, the light produced may be thermal, or coherent. Starting with such a field we would need some phase-sensitive device to produce squeezed light. An analogy with a forced pendulum will illustrate how this can be done.

We start by considering a pendulum formed by a bob hanging at the end of a string supported by a pivot (for our purposes this might be envisioned as a pulley, which will allow us to change the string's length). When the string length is held fixed, the pendulum will undergo simple sinusoidal motion – for sufficiently small amplitude oscillations. If, while the bob is passing underneath the pivot, we temporarily shorten the string, then we shall manage to speed the bob up – just as a figure skater spins faster when pulling in his arms. This will increase the amplitude of the oscillation slightly. Now if we repeat this exercise every time the bob passes underneath – so we pump the pendulum at twice its natural frequency – then we can amplify the amplitude of oscillation. Instead, if we temporarily lower the bob every

time it passes beneath the pivot, we will de-amplify the motion. We can see that we have made a phase-sensitive amplifier! By changing the phase of our pumping the pendulum, we can either amplify or de-amplify its motion.

The electromagnetic field is analogous to a harmonic oscillator, so our discussion of a classical phase-sensitive amplifier gives us the hint as to how we can generate squeezed light from light with phase insensitive noise.

If we somehow pump our signal field at twice its (optical) frequency, then we might expect that not only the average amplitude of the field, but also the fluctuations of the field themselves, will be affected. Indeed, this happens! The quadrature phases represent the different phases of a harmonic oscillator, so such pumping will amplify one quadrature and de-amplify the other. This will produce squeezed light.

Thus, to generate squeezed light, we need to couple our signal modes strongly to some oscillation (even another field) at near twice the signal modes' natural frequencies. To be able to have a strong coupling, which will give contributions even at first order in a perturbation expansion, we must match energies between the pump at $2\omega_0$ and the signal at ω_0. One way to achieve this is to have an interaction Hamiltonian that is quadratic in the signal and linear in the pump:

$$H \propto E_{\text{signal}}^2 E_{\text{pump}} . \tag{I.4.1}$$

This is the kind of interaction that occurs in a nonlinear medium with a second-order nonlinear susceptibility. The resulting device is called a parametric amplifier. Another way to achieve the same result, if the pump frequency is the same as the signal frequency, is to have a Hamiltonian that is quadratic in both the signal and the pump modes:

$$H \propto E_{\text{signal}}^2 E_{\text{pump}}^2 . \tag{I.4.2}$$

This interaction requires a third-order, nonlinear susceptibility, and the resulting device is called a four-wave mixer. We can go on coupling more and more pump modes in, just so long as our energy-matching condition is satisfied (recall that this ensures that the modes couple strongly), but at optical frequencies the required higher order nonlinearities are typically weaker. So far, the best squeezing results have been obtained with a parametric amplifier.[4]

I.5 DIFFICULTIES WITH SQUEEZED LIGHT

In trying to apply the reduced quantum fluctuations in one of the quadrature phases of squeezed light, there are two inherent difficulties. In this section, I will discuss the cause and implications of these difficulties. The first in importance of these difficulties is losses. Any loss in a system carrying squeezed light will by necessity allow ''vacuum'' noise to creep in! This feature is quite general and may be identified as the ''fluctuation-dissipation'' theorem for quantum noise. I will analyze this effect by using a high-transmission beam splitter as a model for the losses. The second problem enters when we try to ''lock on'' to the quadrature with reduced noise. Any fluctuations in the phase, defining which quadrature we are looking at, can in principle cause a small component of the amplified noise to get mixed into our signal. I shall illustrate this problem by considering quantum fluctuations in the local oscillator of the balanced homodyne detector (Sec. I.3).

We start with a look at the effect of losses on squeezed fluctuations. In practice, these losses can arise from many sources: mode mismatching; photodetector inefficiency; mirror transmission; absorption and scattering. For definiteness, however, we shall model these losses via a simple beam splitter (see Fig. 4) with an amplitude coefficient of transmission of $(1-\eta^2)^{1/2}$. A fraction η (η real) of the amplitude

of the incoming signal field E_S is reflected and lost. Along with this loss, however, comes inevitably some light E_N through the ''unused'' port of the beam splitter. The resulting field is

$$E_R = (1-\eta^2)^{1/2} E_S + \eta E_N \ , \tag{I.5.1}$$

and the variance of either quadrature of this field is

$$(\Delta E_{Ri})^2 = (1-\eta^2)(\Delta E_{Si})^2 + \eta^2 (\Delta E_{Ni})^2 \ , \qquad i = 1, 2 \ , \tag{I.5.2}$$

since the signal field E_S and the noise field E_N are assumed to be uncorrelated.

If the signal is squeezed as given in Eq. (I.2.7), and the noise field is simply vacuum, then the variance of the squeezed quadrature becomes

$$(\Delta E_{R2})^2 = \left[\frac{(1-\eta^2)}{G^2} + \eta^2 \right] (\Delta E)^2_{\text{coh}} \ . \tag{I.5.3}$$

Thus, losses introduce phase insensitive fluctuations! These added fluctuations are a problem if $\eta > 1/G$; *i.e.*, if the losses are larger than the squeezing.

One scheme we can use to get around this problem is to ''shine'' squeezed light into the ''unused'' port (with the same quadrature as the signal squeezed), where noise had entered before. If, for instance, we match the squeezing between the noise and the signal, so that

$$(\Delta E_{N2})^2 = \frac{(\Delta E)^2_{\text{coh}}}{G^2} \ , \tag{I.5.4}$$

then we find

$$(\Delta E_{R2})^2 = \frac{(\Delta E)^2_{\text{coh}}}{G^2} \ , \tag{I.5.5}$$

which has the same small fluctuations as the original signal. Crudely speaking, if the losses are small, and the squeezing large, ''shining'' squeezed light at the lossy port (squeezed in the correct quadrature) reduces the losses via

$$\eta \rightarrow \frac{\eta}{G} . \tag{I.5.6}$$

This application of squeezing sounds wonderful. However, its implementation involves two problems:

i) How do we get this squeezed light for the unused port? We can't just use a beam splitter on what we have in order to make two beams of squeezed light; we have just shown what a beam splitter does to the squeezing. Instead, we must take some unsqueezed light and squeeze it for each and every use we have in mind for the squeezing.

ii) Not all losses are quite so easy to control as our beam-splitter model suggests. For instance, losses caused by absorption or scattering open up too many and too complicated channels for our ''loss squeezing'' method to work. In fact, it appears that for the best mirrors available today (as used in laser gyroscopes, having energy losses of one part in 10^5), perhaps one-half of these losses are due to absorption and scattering.[5] It is exactly these uncontrollable losses that limit the utility of squeezed light in experiments requiring their reduced noise.

The second difficulty with using the reduced noise in squeezed light is the problem of being able to pick out one specific quadrature – either to look at, or to try to squeeze. I shall illustrate the problem through the way it affects our analysis of the homodyne detector (Sec. I.3). We want to see the effects of fluctuations we neglected in Eq. (I.3.3). To do this really requires quantum mechanics, but we can get a good idea from heuristic classical arguments. The main feature I want to include is the

fluctuation of the local oscillator – once all other sources of noise are dealt with, we are still left with the local oscillator quantum noise. Since the phase of the local oscillator determines which quadrature is picked out [Eq. (I.3.6)], any phase noise in the local oscillator will tend to mix in some of the quadrature we are not interested in.

Suppose we are interested in looking at only the squeezed quadrature E_{S2} of some signal field [squeezed according to Eq. (I.2.7)]. Ideally, we want $\theta = \pi/2$ in Eq. (I.3.6), which would give

$$I_D \propto E_{S2} \tag{I.5.7}$$

and

$$(\Delta I_D)^2 \propto (\Delta E_{S2})^2 . \tag{I.5.8}$$

However, because of the phase noise ($\Delta\phi$) of the local oscillator, we find that

$$(\Delta I_D)^2 \propto (\Delta E_{S2})^2 + (\Delta E_{S1})^2 (\Delta\phi)^2 . \tag{I.5.9}$$

If the local oscillator is in a coherent state, then its phase uncertainty is

$$(\Delta\phi) = \frac{1}{2\sqrt{N}} , \tag{I.5.10}$$

where N is the number of photons in the local oscillator within a wave packet whose length is determined by some bandwidth for the system. This makes the contribution to the variance of the homodyne detector's difference current

$$(\Delta I_D)^2 \propto (\Delta E_{S2})^2 + \frac{(\Delta E_{S1})^2}{4N} . \tag{I.5.11}$$

Thus, when we try to use the homodyne detector to look at the squeezed quadrature,

we inevitably get some contributions from the amplified quadrature (clearly, the converse process is not a problem).

When is this phase noise a problem? Clearly, we can get no further improvement by increased squeezing when the two terms on the right-hand side of Eq. (I.5.11) are comparable. For ideal squeezed light we may relate the noises in the two quadrature phases by Eq. (I.2.7). In this case, Eq. (I.5.11) shows that the local oscillator phase noise becomes a limiting factor when

$$N \lesssim \frac{G^4}{4} . \tag{I.5.12}$$

The issue I haven't addressed in this analysis yet is: What is the bandwidth that determines how many quanta are in the local oscillator? For a local oscillator with power P and "bandwidth" B, the number of quanta is

$$N = \frac{P}{B\,\hbar\omega_o} . \tag{I.5.13}$$

Surprisingly, the relevant bandwidth is the bandwidth over which the signal is squeezed. If it had been the bandwidth of the vacuum fluctuations around the local oscillator, then the phase-noise problem would be arbitrarily large. As I have stated, the justification for the choice of this bandwidth requires quantum mechanics. For a laser of 1 mW at optical frequencies, being used as the local oscillator for a homodyne detector of broad band squeezing having a bandwidth $B \sim 10^{10}\,\mathrm{sec}^{-1}$, the number of photons is

$$N \sim 10^6 . \tag{I.5.14}$$

In this case we couldn't see better than a factor of $G^2 \sim 1000$ squeezing in noise power

in the output of the homodyne detector. This is not a serious limit yet. It should be noted that the current best[4] attempts at squeezing give $G^2 \sim 3$.

This second problem of defining the phase accurately also occurs when we try to make squeezed light. David Crouch and I have studied this problem, when the device used to make the squeezed light is a parametric amplifier, as discussed in Chapter III.

I.6 IMPROVEMENT IN AN INTERFEROMETER

I now want to give a very brief discussion of the pioneering work of Caves[6] on the improvement of the sensitivity of an interferometer by using squeezed states of light. It is somewhat surprising that the limit of phase sensitivity that is due to counting statistics (shot noise) in an interferometer can be improved by "shining" squeezed light into the interferometer's output port!

Let us consider a very simple Michelson interferometer. The light is incident on a 50–50 beam splitter, where it is directed down either of a pair of arms ending with perfect reflectors. After the light returns to the beam splitter it is recombined there. The 50–50 beam splitter has four ports: two external ones (to the interferometer) labeled A and B, which serve as inputs to and outputs from the interferometer, and two internal ones, C and D, which lead to and return from the arms of the interferometer. We will label the arms of the interferometer so that external port A (B) of the beam splitter looks straight down (forgetting reflections) arm C (D). Suppose we illuminate port A with a laser (with field E_L). We shall be interested in observing the interference between the beams of light that travel down the two arms and recombine as the output of port B of the beam splitter.

As I have mentioned in Sec. I.5, a beam splitter allows noise from input port B (with field E_N) into the apparatus. So our analysis must include both input fields E_L

and E_N to see how using squeezed light might help. Just after these fields enter the interferometer from the outside, the field amplitudes at the start of arms C and D will be

$$E_A = \frac{1}{\sqrt{2}}(E_N + E_L)$$

$$E_B = \frac{1}{\sqrt{2}}(E_N - E_L)\,. \tag{I.6.1}$$

After these combined fields have traveled down the arms and returned, but before they pass through the beam splitter, they will have acquired a phase difference ϕ that is due to the path difference in the two arms. We may write this as

$$E_A = \frac{1}{\sqrt{2}}(E_N + E_L)\,e^{i\phi/2}$$

$$E_B = \frac{1}{\sqrt{2}}(E_N - E_L)\,e^{-i\phi/2}\,. \tag{I.6.2}$$

The resulting field amplitude observed at output port B is then

$$E_R = iE_L \sin(\phi/2) + E_N \cos(\phi/2)\,. \tag{I.6.3}$$

We see that the laser field has acquired a phase shift of 90° because of its coming out of a different port from the one it entered. For small changes in phase around a null for the laser at the output port B (*i.e.*, $\phi \sim 0$), we have

$$E_R \sim \frac{i}{2}E_L\phi + E_N\,. \tag{I.6.4}$$

This is almost the entire story for the phase sensitivity of an interferometer. The amplitude of the laser $<E_L>$ is being used as a large lever arm to magnify the size of

the phase changes in the interferometer; this is described by the first term in Eq. (I.6.4). Thus, any noise that causes fluctuations inphase with the amplitude of the laser will limit the sensitivity of the interferometer. As the laser field has acquired a 90° phase shift relative to the noise field [see Eq. (I.6.4)], this translates into phase fluctuations (relative to the laser's mean-field phase) for the noise field E_N. That is, it is the ''phase'' fluctuations in the noise field that are responsible for the limit to the phase sensitivity. If this noise field is in a vacuum state, then the phase sensitivity is given by the ''shot-noise'' limit

$$(\Delta\phi) \sim \frac{1}{\sqrt{N}} , \tag{I.6.5}$$

which corresponds to counting statistics for N independent photons. If, instead, we squeeze the ''phase'' quadrature of the field entering the unused port of the interferometer, say according to Eq. (I.2.7), then we find

$$(\Delta\phi) \sim \frac{1}{G\sqrt{N}} . \tag{I.6.6}$$

Amazingly, this kind of improvement in the sensitivity of an interferometer has been achieved experimentally. Xiao et al.[7] found that squeezing could be used to produce a reduction of the noise in an interferometer by 3.0dB below the shot-noise limit.

I have already discussed in general terms in Sec. I.5 what can limit the utility of squeezed light; the same principles apply for using squeezed light in an interferometer.

I.7 QUANTUM VARIABLES

In this section, I introduce some definitions that shall be used extensively in Chapter III. The main concern is to simplify a fully quantum treatment by working directly with the mode annihilation and creation operators which are the Fourier components of the electromagnetic field. However, since we have found the quadrature phases such a useful tool in the time domain, we seek similar objects in the frequency domain. We do this by expanding the electric field into mode annihilation and creation operators and then relating them to the Fourier components of the quadrature phases. Fourier components of these quadrature phases are called the quadrature-phase amplitudes.

Almost our entire discussion so far has used language that is very well suited to thinking of the fields in the time domain, namely,

$$E(t)=E_1(t)\cos(\omega_0 t)+E_2(t)\sin(\omega_0 t)\,. \tag{I.7.1}$$

However, when dealing with quantum systems, it is useful to move to the frequency domain and to decompose the resulting fields into annihilation and creation operators, *i.e.*, the positive and negative frequency components of the field. That is, we wish to decompose the electric field as

$$E(t)=E^{(+)}(t)+E^{(-)}(t)\,, \tag{I.7.2}$$

where

$$E^{(+)}\propto\int\frac{d\omega}{2\pi}\sqrt{\omega}\,a(\omega)e^{-i\omega t}\,,\qquad E^{(-)}=E^{(+)\dagger}\,. \tag{I.7.3}$$

Writing the electric field in terms of the quadrature-phase amplitudes requires "demodulating" this field at frequency ω_0, as in Eq. (I.7.1). This causes the spectrum

of E_1 and E_2 to be that of the E field but folded over at ω_0 (the fold becoming labeled zero frequency). A little algebra yields

$$X_1(\varepsilon) \equiv \int \frac{dt}{2\pi} e^{i\varepsilon t} E_1(t) = E^{(+)}(\omega_0+\varepsilon) + E^{(-)}(\omega_0-\varepsilon) \tag{I.7.4}$$

$$X_2(\varepsilon) \equiv \int \frac{dt}{2\pi} e^{i\varepsilon t} E_2(t) = -iE^{(+)}(\omega_0+\varepsilon) + iE^{(-)}(\omega_0-\varepsilon)\,. \tag{I.7.5}$$

These new variables are called the quadrature-phase amplitudes. For small bandwidths $\varepsilon \ll \omega_0$, we may write

$$X_1(\varepsilon) = \frac{1}{2}[a(\omega_0+\varepsilon) + a^\dagger(\omega_0-\varepsilon)] \tag{I.7.6}$$

$$X_2(\varepsilon) = \frac{i}{2}[a^\dagger(\omega_0-\varepsilon) - a(\omega_0+\varepsilon)]\,, \tag{I.7.7}$$

which are frequency components of the quadrature phases directly in terms of annihilation and creation operators. These new variables carry directly the spectral information that is seen in the power spectrum of an optical homodyne detector. Except at $\varepsilon = 0$, these variables are not Hermitian and so are not directly observable. Their real and imaginary parts, however, are Hermitian and form the observables seen directly in the difference current of a balanced homodyne detector.

I.8 INTRODUCTION TO CHAPTER II

In Chapter II, Robert McLachlan and I study the states generated by a generalized parametric amplifier. This generalized device produces k-photon correlations by pumping a signal mode at k times its natural frequency. For $k = 2$, this device will generate squeezed light (as argued in Section I.4). In this case, the pump is in-phase with the signal at two points around a phase-space diagram, and so the amplification

produces an ellipse [Fig. 1(c) and (e)]. The k-photon parametric amplifiers with $k > 2$ are considered to produce a generalization of a squeezed state. Since now the signal will be inphase with the pump at k angles around the phase-space diagram, we expect the phase-space diagram to look like a k-lobed star. Thus, these generalized squeezed states will have at least one important property in common with ordinary squeezed light: phase sensitive noise.

This k-photon parametric amplifier had been considered before by Fisher et al.,[8] who looked at the Taylor series expansion of the vacuum expectation of the evolution operator of these devices. They found that this expansion had zero radius of convergence for $k > 2$, which they interpreted as showing that the model for these devices was meaningless. There was no apparent physical reason for this failure of the model, so that led to suspicion of their results.

It is well known that power series that are only asymptotically convergent may be expansions of perfectly reasonable functions. This led us to hope that the expansion of the fatal matrix element was an asymptotic series. We tried applying various techniques to improve the convergence of this series. Because of the complicated nature of the coefficients, we did this numerically. Robert McLachlan got the first finite results using the method of Padé approximants,[9] which we continued to use for all our numerical work on these models.

We found that the region over which we could obtain convergence was limited only by the accuracy to which we could perform the numerical calculations. We also showed that every matrix element of these models could be obtained from differentiation of the matrix element studied by Fisher et al.[8] We were even able to investigate the pole structure of this matrix element, using our numerical methods. We found that there appeared to be an accumulation of poles near the origin (along the imaginary

''time'' axis). As we increased the accuracy of our calculations, more of these poles appeared (see Fig. 3 of Chapter II). This strongly indicated that the origin was a branch point; this would explain why the Taylor series of the matrix element had a zero radius of convergence about this point. We concluded from this that the models were almost certainly well defined.

Strictly speaking, phase space doesn't exist for quantum systems: a point on a phase-space diagram defines a point with an exactly specified pair of canonically conjugate variables, which would violate the Heisenberg uncertainty principle. It is, however, possible to define something analogous. For instance, we may take the expectation value of the quantum density matrix in a harmonic oscillator coherent state. This function is a well-defined probability distribution and is called the quantum Q-function.[10] It measures the probability of the quantum state being found in a coherent state of complex amplitude α. Plotting this Q-function as a function of the real and imaginary parts of α then gives a diagram that is essentially a ''quantum phase space picture.''

We used our numerical results to plot the Q-function for various times in the evolution of the states generated from vacuum for these devices with $k = 3$ and $k = 4$ [see Figs. 1.(c)-(f) in Chapter II]. These matched our expectation of k-lobed stars, and we compared these to calculations of the corresponding classical models for these devices. The major difference between the classical and quantum evolution was an extra ''quantum diffusion,'' which prevented the Q-function from developing very sharp structures (and hence from violating the Heisenberg uncertainty principle).

Finally, in an analytic calculation to second order in time, we found that an initial coherent state displaced from the origin (along one of the axes where a lobe would have formed from a vacuum state) develops ordinary squeezing. This result was valid

for any of the devices with $k \geq 2$ and might be useful in the future for generating ordinary squeezed light (this result had been noted before).[11]

I.8 INTRODUCTION TO CHAPTER III

In Chapter III, David Crouch and I consider the limit to squeezing in a parametric amplifier with a quantum pump. Prior to this work there had been three other calculations[12–14] along similar lines with different regions of validity, all on slightly different models. Only two of these calculations[12,14] agree with each other; however, the regions of validity for these two calculations do not extend to the point where their limitation becomes relevant. In all these calculations, results were obtained only to $O(1/\bar{N})$, where $\sqrt{\bar{N}}$ is the amplitude for the pump, assumed to be in a coherent state.

We derive results for each of the models to $O(1/\bar{N}^2)$ in order to check the prior results, and where necessary we extend the regions of validity to the point where a limit directly manifests itself. Because of the disagreement among the many previous calculations, we have felt justified in using more than one method. Some of the chapter can be neatly separated into parts done solely by one or another of the authors: Section 6 (unlabeled sections here refer to sections in Chapter III) is David's work, as is the argument for obtaining discrete mode equations for a continuum system (which appears as part of Section 2); Sections 3, 4 and 5 are, however, my own work. The remainder of the chapter should be considered as joint work; the chapter has been left intact with the contributions from both authors to preserve its exposition.

The central model studied in Chapter three is the multimode parametric amplifier: many signal modes coupled to a single pump mode. This model can cover any of the previous models by fixing the number of signal modes appropriately.

In Section 3, we derive the symmetries of the equations of motion of this multimode parametric amplifier and, using them, we shorten the list of undetermined matrix elements relevant to questions about the squeezing achievable in the model.

The first method of calculation (Section 4) involves integrating the Heisenberg equations in terms of the quadrature phase variables up to $O(1/\bar{N}^2)$. What is novel about the approach here is that we first identify which terms can contribute to the dominant effects, and then we proceed to calculate only those terms. We thus derive the dominant terms up to $O(1/\bar{N}^2)$. By calculating those terms, we could estimate the time when the $O(1/\bar{N})$ correction (the one yielding the limit to squeezing) breaks down. We find that this $O(1/\bar{N})$, *i.e.*, semiclassical, correction is valid so long as it is much less than one.

The second and third methods restrict their efforts to the one- and two-mode parametric amplifiers.

The second method (Section 5) is numerical, but we are able to obtain analytic expressions for the full corrections to $O(1/\bar{N}^2)$, which if performed solely via standard analytic methods would have required thousands of terms at the intermediate stages. The conversion from numerical to analytic is possible because the general form of the corrections was already determined in Section 4. At the heart of the calculation is a new algebra, which can be viewed as a semiclassical (or higher-order) approximation to the ordinary commutator algebra for annihilation and creation operators.

Finally, the third method (Section 6, by David Crouch) uses the positive-P distribution[15] to derive Fokker-Planck equations. The positive-P distribution is analogous to the Q-function used in Chapter II and is a generalized quantum phase space distribution. Standard methods of stochastic calculus[16] are then used to derive stochastic differential equations. Approximate solutions to these stochastic equations were then

obtained by iteration to obtain the $O(1/\bar{N})$ corrections only. This method is equivalent to a direct iteration of the operator Heisenberg equations of motion but in the language of stochastic differential equations.

The basic result of the paper is that the initial phase noise in the pump causes an uncertainty in which axes (in a phase-space diagram) will be chosen to be squeezed. Whatever axis is picked out on average, there will be a contribution from the orthogonal (the amplified) quadrature as the pump's phase fluctuates. When the pump is initially in a coherent state, this limits the variance of the squeezed quadrature to $O(1/\sqrt{\bar{N}})$.

II.1 INTRODUCTION TO PART II

In part II of this introduction, I will start in Section 2 with a brief description of measurements in quantum mechanics. I follow this with an introduction to Chapter IV. In Section 3, I review what I consider to be the status of the problem of objectivity in quantum mechanics: Where does the objectivity lie? Section 3 ends with a discussion of the work and results of Chapter V.

Even accepting, without question, the correctness of the Schrödinger equation, there are areas of the interpretation and usage of quantum mechanics that are perhaps still seen as problematic. Just some of these areas are collapse of the wave function and its associated "spooky actions at a distance," and the question of whether objectivity can be consistent with a quantum description. Although these sorts of issues seem to be purely interpretational, it is becoming apparent that the foundations of quantum mechanics can lead directly to testable propositions.

II.2 INTRODUCTION TO CHAPTER IV

Measurements in quantum mechanics can be described by a "measurement model": the system is coupled to a "measuring apparatus;" after the interaction, one obtains information about the system by observing some property of the apparatus. It is advantageous to formalize this model by the introduction of mathematical objects called Effects and Operations.[17] A set of Effects generates from the system state the "detector statistics," *i.e.*, the statistics of the possible outcomes. Once a particular result has been obtained from the measurement, an Operation maps the system state before the measurement to the new (unnormalized) system state after the measurement. These objects are generalizations of elementary concepts: the probability of finding a system in some specific quantum state (given by the absolute value squared of the inner product between the system state and the state of interest); and the "collapse" of the wave function when the system is found to be in some specific state.

Recently, a paper by d'Espagnat[18] has brought up questions about the usage of this formalism by Barchielli, Lanz, and Prosperi.[19] The latter authors used Effects and Operations to develop a formal description of a continuous measurement of position. To generate the measurement statistics, they used a simple set of Effects. Further, they chose the simplest Operations consistent with their Effects (*i.e.*, their measurement statistics). Even a careful reading of their paper might lead one to believe that they are proposing their particular choice of Operations as a fundamental constituent in a new theory of measurement. Only in their Appendix B do they make it clear that their choice of Operations is not unique; it corresponds to some measurement model that they do not specify.

D'Espagnat[18] takes the Operation of Barchielli *et al.* as part of a new theory, which he dubs the "Milano theory." He then compares this theory with standard

quantum mechanics by analyzing two consecutive measurements of spin by two different methods. In the first method, he formulates a measurement model and applies to it the standard quantum rules for calculating probabilities. In the second method, he applies to two measurements of spin an Operation analogous to the one of Barchielli *et al.* Not surprisingly, he finds that the two methods disagree. The Operation of the "Milano theory" corresponds to *some* model for measurements of spin, but it is not the measurement model that d'Espagnat has formulated.

In Chapter IV, Carlton Caves and I explore and explain the discrepancy found by d'Espagnat. We start by reviewing briefly the formalism of Effects and Operations and show how it arises naturally from applying the standard quantum rules to a measurement model. We study d'Espagnat's model in detail and discuss how it is handled within the framework of Effects and Operations. Finally, we consider the Operation for a spin measurement from the "Milano theory," and we formulate a measurement model, different from d'Espagnat's, that realizes this Operation. This shows explicitly that d'Espagnat was comparing two different models.

II.3 INTRODUCTION TO CHAPTER V

What is the problem?

For those people who think that the wave function is an objective quantity, collapse of the wave function poses a big problem. Why should an objective wave function change at all when we obtain information? Where does the rest of the wave function "go" when we make a measurement? Perhaps the most logical solution to this problem is the "many-worlds" interpretation[20] of quantum mechanics. Unfortunately, this interpretation retrieves the rest of the wave function only to proceed to place it in other branches of the universe where it remains totally inaccessible forever

from our own branch. This interpretation, then, leaves us with an untestable proposition.

If, instead, one regards the wave function as a mathematical tool – a repository of information – then there is no problem with wave function collapse. The standard quantum rules tell us that when the value of some observable is learned, we should throw away those parts of the wave function that are inconsistent with our new knowledge. This is analogous to a probabilistic description of a classical system in which the system has objective properties, but the probability is not itself objective; if we learn something about the system, we reduce the probability conditioned on that information (this reduction of probabilities is just Bayes's theorem for conditioned probabilities). This simple solution holds, even in a general description of quantum measurements – such as the formalism of Effects and Operations (reviewed in Section IV). The problem now, however, is whether or not the information stored by the wave function is information about objective properties. And further, if these properties described by the wave function are objective, why can't we use ordinary probability logic for them, instead of being forced to use quantum mechanical amplitude logic?

I see the deeper problem as this: how we can get objective properties – "what we actually see in experiments" – when quantum mechanics is manifestly not objective. This is sounding like a discussion on quantum theology – "why do we see the world as classical and objective?" I won't attempt to answer this question but instead want to show that quantum mechanics – or any theory that agrees with quantum statistics – cannot be viewed as an objective theory. Such a more complete theory with an objective interpretation was Einstein's program.

Let us start by reviewing Bohmand Aharonov's[21] version of the Einstein-Podolsky-Rosen (EPR) paradox.[22] For a pair of spin one-half particles[21] formed in a

spin-singlet state, the EPR argument may be stated thus. Measurements of the components of the spin $\vec{\sigma}_1$ and $\vec{\sigma}_2$ for each particle can be made. If a measurement of $\vec{\sigma}_1 \cdot \vec{b}$ (where $\vec{b}$ is a unit vector along the orientation of a detector) yields +1, then a subsequent measurement of $\vec{\sigma}_2 \cdot \vec{b}$ will yield −1. Now if the detectors are separated by a great distance so that they cannot communicate (the locality assumption) during these measurements, then the result of $\vec{\sigma}_2 \cdot \vec{b}$ must be predetermined, since we could have predicted it from a measurement of $\vec{\sigma}_1 \cdot \vec{b}$. But its being predetermined should not depend on our first measuring $\vec{\sigma}_1 \cdot \vec{b}$, so it must be a "concrete," objective property of particle 2. The properties of objectivity and locality are called *local realism.*

This argument can be repeated for a different direction $\vec{b}'$ for particle 2 with the same conclusion; $\vec{\sigma}_2 \cdot \vec{b}'$ would be a realistic property . Similarly, any sets of directions $\vec{a}, \vec{a}', \cdots$ for particle 1 would have predetermined results for $\vec{\sigma}_1 \cdot \vec{a}, \vec{\sigma}_1 \cdot \vec{a}', \cdots$, each being realistic properties.

Now quantum mechanics with its description of the state via a wave function does allow predictions for pairs of these spin projections, but it has no notion of more than one pair at a time – these properties don't commute, so we can't talk about them quantum-mechanically. Thus, this large set of realistic properties leads us to conclude that quantum mechanics is an incomplete description of reality – that somehow there is much more information stored in the real physical system than quantum mechanics allows for.

Although the EPR argument is at first sight compelling, Bell[23] took this concept of objectivity and derived an inequality that correlation functions from any realistic theory must satisfy – thus making the question of objectivity testable. Various versions of this inequality are violated by quantum mechanics. These inequalities apply to pairs of two-level systems just as in our example of the EPR paradox. We shall

present a simple proof of one of these inequalities.[24]

If spin components are objective quantities, then we can write down expressions for each individual decay. Consider, for instance, the expression

$$\sigma_a\,(\sigma_b+\sigma_{b'})+\sigma_{a'}(\sigma_b-\sigma_{b'})\equiv\pm 2\,, \tag{II.3.1}$$

where σ_a and $\sigma_{a'}$ are the spin projections of particle 1 along axes $\vec{a}$ and $\vec{a}'$, respectively, and similarly, σ_b and $\sigma_{b'}$ are the projections of particle 2 along axes $\vec{b}$ and $\vec{b}'$. This equality is trivially true for each decay, since each spin projection is ± 1, and since for each decay the results σ_a, $\sigma_{a'}$, σ_b and $\sigma_{b'}$ are all predetermined, and as such are well defined. Over many decays these objective properties could, in principle, be accumulated to form a grand probability distribution

$$p\,(\sigma_a\,,\sigma_{a'},\sigma_b\,,\sigma_{b'})\,. \tag{II.3.2}$$

In practice, we would measure only pairs of these spin projections for each decay; then after we had measured these pairs for all detector ''settings'' (two for each side), we could incorporate what information we had learned into this grand probability.

This grand probability of Eq. (II.3.2) may now be used to average Eqn. (II.3.1) over many decays. This averaging gives the Clauser-Horne-Shimony-Holt[25] (CHSH) Bell inequality.

$$|C\,(\vec{a},\vec{b})+C\,(\vec{a},\vec{b}')+C\,(\vec{a}',\vec{b})-C\,(\vec{a}',\vec{b}')|\leq 2\,; \tag{II.3.3}$$

here the correlation function is just the expectation of the product of the spin projections

$$C\,(\vec{a},\vec{b})\equiv<\sigma_a\,\sigma_b>\,. \tag{II.3.4}$$

A brief review of the experimental status of this inequality may be found in Chapter V.

Carlton Caves and I further investigate the consequences of objectivity in Chapter V. We start by introducing the auxiliary concept of chaining the CHSH Bell inequality. Although this chaining should apparently lead to a "weaker" condition for local realism, these new chained correlation Bell inequalities actually have stronger quantum violations (*i.e.*, larger signal-to-noise ratio) over a larger set of angles. To compute a signal-to-noise ratio, we took a simple model of an experimental setup: Counting statistics determine the variances of the measured correlation functions, and a "data flipping error" (*i.e.*, the detector channels – the "±1" – can get confused) degrades the ideal quantum-mechanical correlation function. When the data error quantity is matched to that of the most recent experiments,[26] we expect a 20 % improvement in signal-to-noise ratio using our first chained correlation Bell inequality compared to using the CHSH Bell inequality. As this data error increases, the advantage of chaining decreases, the more heavily chained inequalities being most severely affected.

Next, we take seriously the idea that systems with objective properties would carry more information than their quantum descriptions seem to allow. We do this by deriving new constraints on local realism in terms of the average information (analogous to entropy) obtained in making a measurement. These new inequalities, which we call information Bell inequalities, are easy to derive and apply to many new situations, since they rely only on the existence of a grand probability, like that which appears in Eqn. (II.3.2), and a locality assumption. As an example, we derived information inequalities for higher-spin versions of the EPR experiment. We find that these inequalities are violated by quantum mechanics for arbitrarily high values of the

spin and further, that the size of the violation increases with increasing spin. The size of the violation for information Bell inequalities has a direct interpretation as the amount of information by which the quantum description is deficit compared to even the most thrifty, locally realistic description. The chaining concept can also be applied to these information Bell inequalities, and will yield improvements similar to those found for the CHSH correlation Bell inequalities.

Finally, we note that the generality of the assumptions of the information Bell inequalities leads to a natural hierarchy of Bell inequalities for many particle systems. Even in this general hierarchy, however, the two-particle Bell inequalities play a special role in that they lead to the simplest direct tests of local realism.

REFERENCES

[1] C. M. Caves and B. L. Schumaker, Phys. Rev. A **31**, 3068 (1985).

[2] R. J. Glauber, Phys. Rev. **130**, 2529 (1963).

[3] H. P. Yuen and J. H. Shapiro, IEEE Trans. Inf. Theory **24**, 657 (1978).

[4] L.-A. Wu, H. J. Kimble, J. L. Hall and H. Wu, Phys. Rev. Lett. **57**, 2520 (1986).

[5] Private communication from Al Bostik.

[6] C. M. Caves, Phys. Rev. D **23**, 1693 (1981).

[7] M. Xiao, L-A. Wu and H. J. Kimble, Phys. Rev. Lett. **59**, 278 (1987).

[8] R. A. Fisher, M. M. Nieto and V. D. Sandberg, Phys. Rev. D **29**, 1107 (1984).

[9] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1978), p.383.

[10] K. E. Cahill and R. J. Glauber, Phys. Rev. **177**, 1882 (1969).

[11] M. Hillery, M. S. Zubairy and K. Wódkiewicz, Phys. Lett. **103*A***, 259 (1984).

[12] M. Hillery and M. S. Zubairy, Phys. Rev. A **29**, 1275 (1984).

[13] G. Scharf and D. F. Walls, Opt. Commun. **50**, 245 (1984).

[14] C. M. Caves and D. D. Crouch, J. Opt. Soc. Am. B **4**, 1535 (1987).

[15] P. D. Drummond and C. W. Gardiner, J. Phys. A **13**, 2353 (1980).

[16] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1983).

[17] K. Kraus, *States, Effects, and Operations: Fundamental Notions of Quantum Theory* (Springer-Verlag, Berlin, 1983).

[18] B. d'Espagnat, Found. Phys. **16**, 351 (1986).

[19] A. Barchielli, L. Lanz and G. M. Prosperi, Nuovo Cimento B **72**, 79 (1982).

[20] Everitt, *Many Worlds Interpretation of Quantum Mechanics,* (Princeton University Press, New Jersey, 1980).

[21] D. Bohm and Y. Aharonov, Phys. Rev. **108**, 1070 (1957).

[22] A. Einstein, N. Rosen and B. Podolsky, Phys. Rev. **47**, 777 (1935).

[23] J. Bell, Physics **1**, 195 (1964).

[24] A. Peres, Am. J. Phys. **46**, 745 (1978).

[25] J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[26] A. Aspect, P. Granier and G. Roger, Phys. Rev. Lett. **49**,91 (1982).

FIGURE CAPTIONS

FIG. 1. Plot of the quadrature phases on a phase-space diagram with a nonzero component of the average field in E_1, and fluctuations for E_1 and E_2 displayed as a circle or ellipse, for (a) a coherent state, (c) a phase-squeezed state, and (e) an amplitude-squeezed state. The appropriate time domain plot of the electric field is shown for each of the cases (a), (c) and (e) in (b), (d) and (f), respectively. [Figure taken from Caves[6] with permission].

FIG. 2. Schematic of a balanced homodyne detector. A signal field E_S and local oscillator E_{LO} impinge on the 50–50 beam splitter from different directions. The summed and differenced amplitudes made in this way are incident on a pair of photodetectors. The photodetectors each produce a photocurrent proportional to the square of the electric field amplitude falling on them. The difference I_D of these currents forms the output of the homodyne detector. [Figure taken from Caves[6] with permission].

FIG. 3. Plot of the rms noise voltage – *i.e.*, the square root of the noise power – versus the local oscillator phase for the fixed RF frequency of 1.8 MHz over a bandwidth of 100 kHz, for the experiment of Wu et al.[4] The dashed horizontal line is the vacuum-noise level. [Figure taken from Wu[4] with permission].

FIG. 4. A simple beam splitter with which we model losses. A signal E_S impinges from the left, and a "noise field" E_N from below. A nonzero reflection coefficient allows most of E_S through to the right; however, a portion of E_N accompanies it.

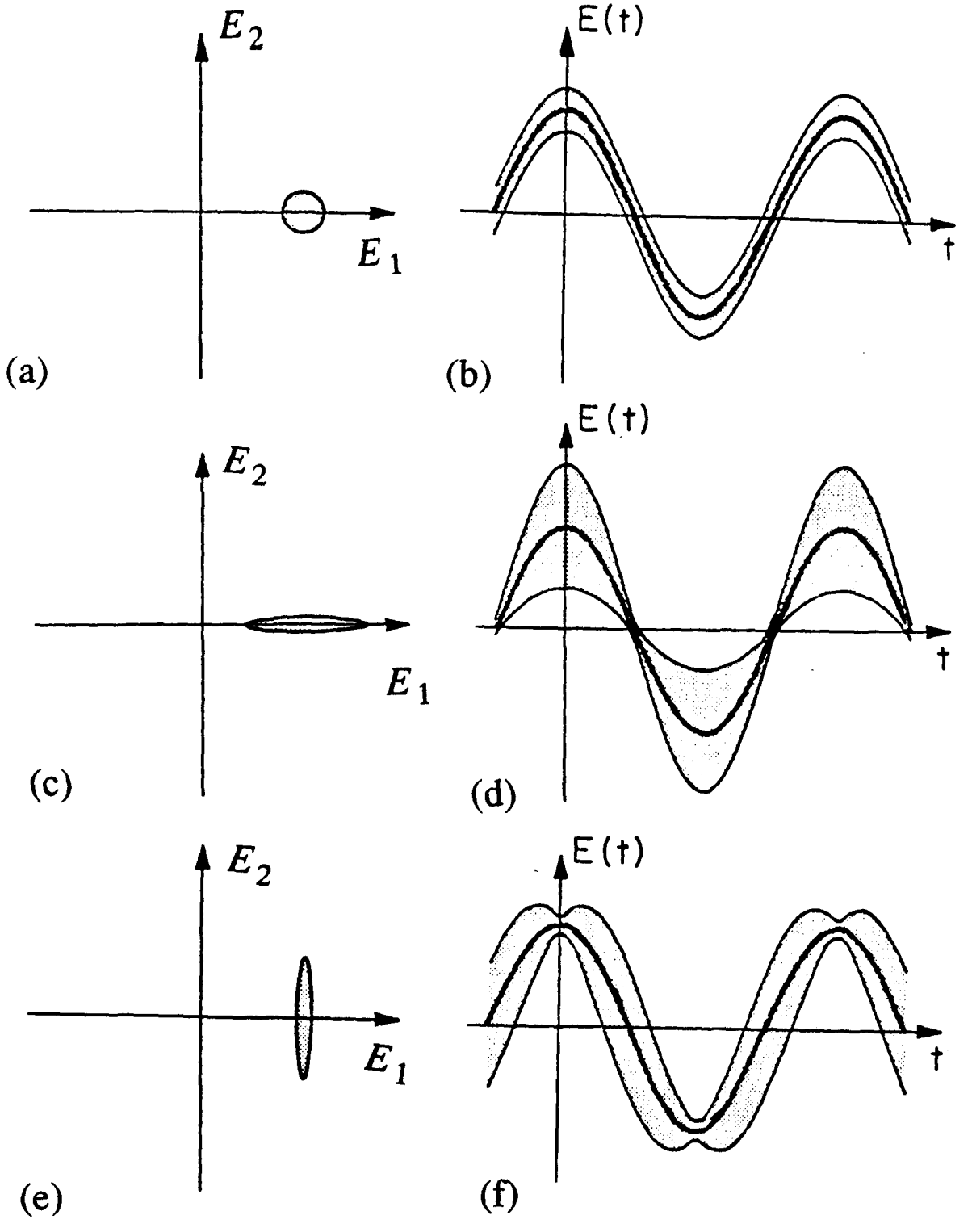

Figure 1

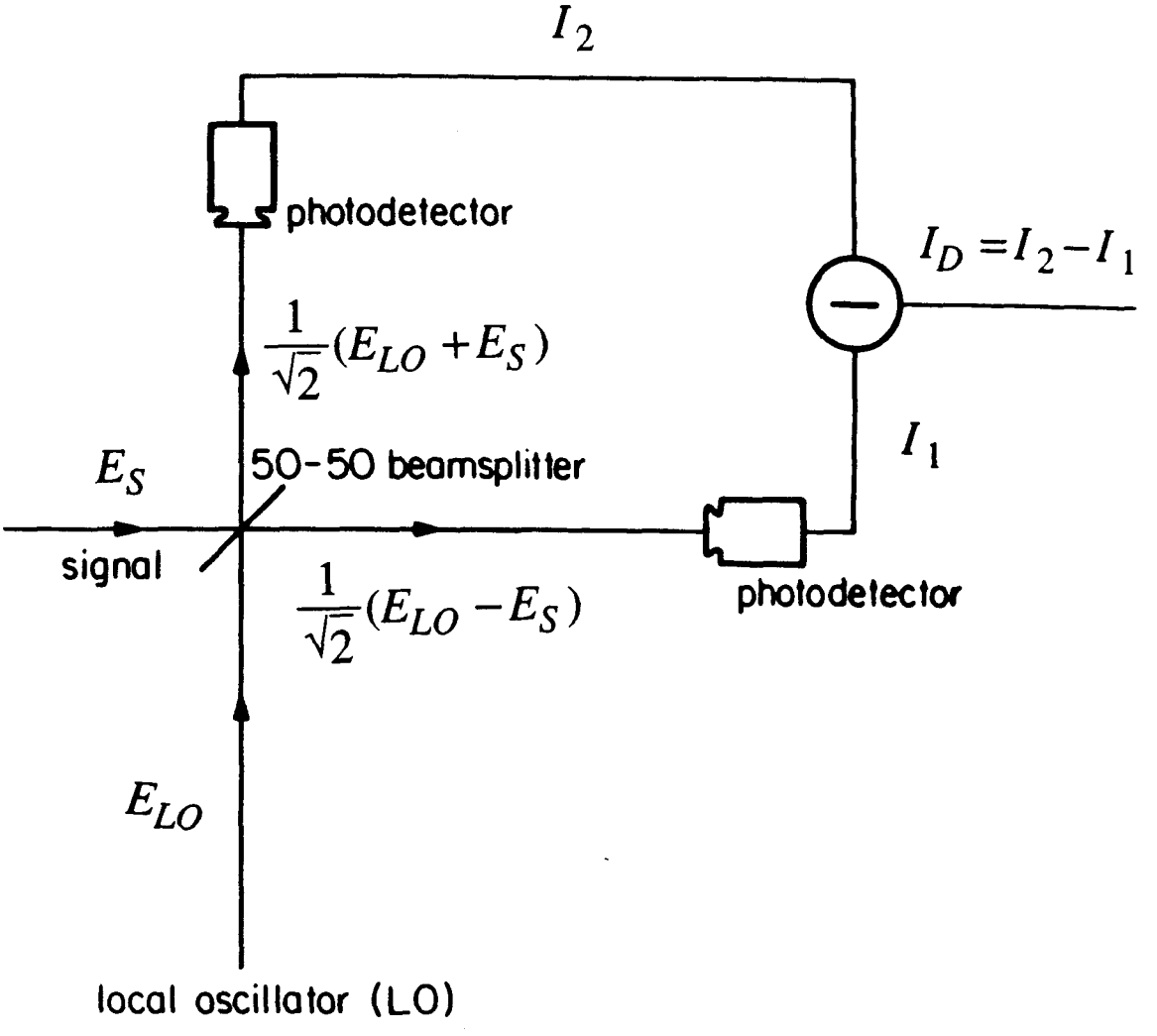
I_2
photodetector
$I_D = I_2 - I_1$
$\frac{1}{\sqrt{2}}(E_{LO} + E_S)$
I_1
E_S
50-50 beamsplitter
signal
$\frac{1}{\sqrt{2}}(E_{LO} - E_S)$
photodetector
E_{LO}
local oscillator (LO)

Figure 2

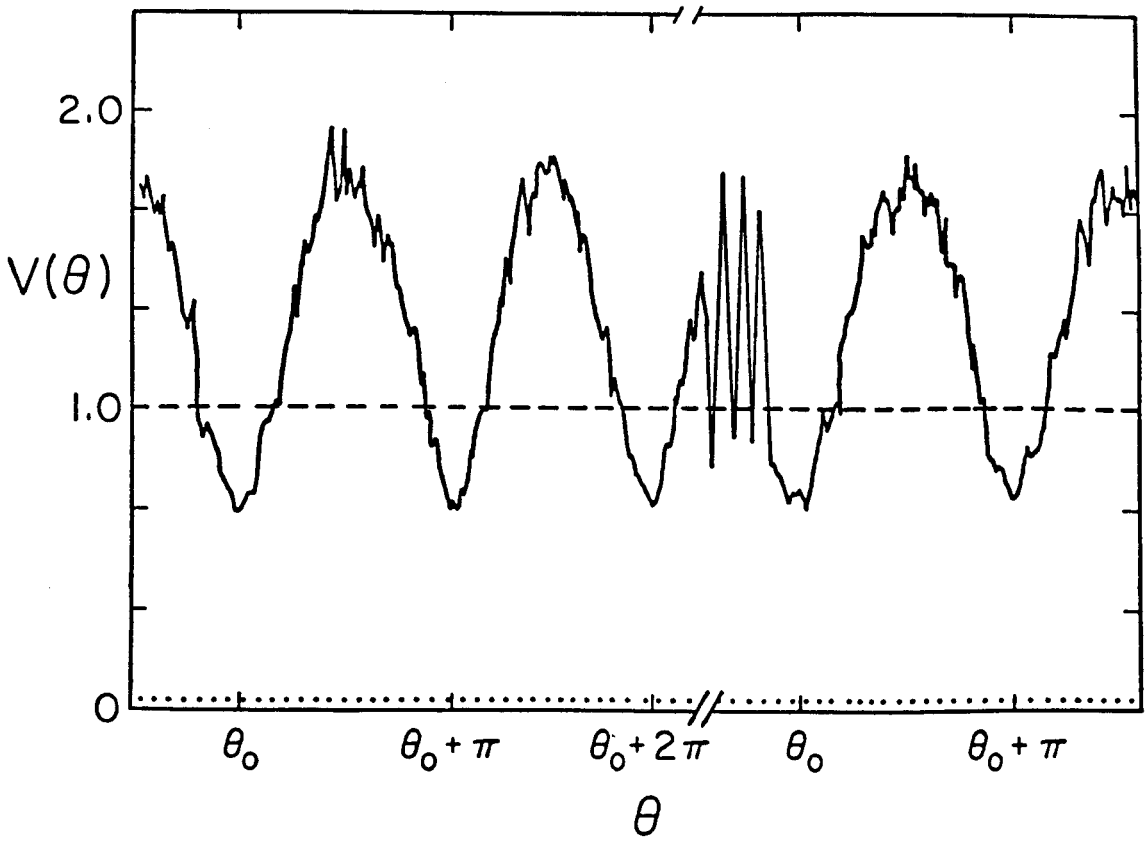

Figure 3

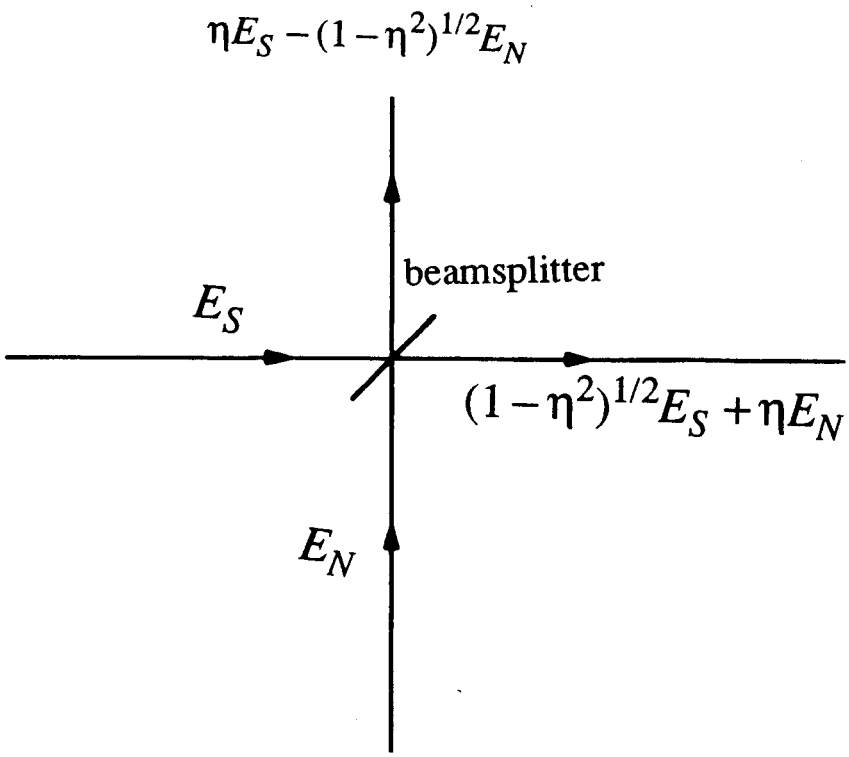

Figure 4

PART ONE

CHAPTER 2

Generalized squeezing

by Samuel L. Braunstein and Robert I. McLachlan

Published in Physical Reveiw A.

ABSTRACT

We consider a generalized form of parametric amplification that produces k-photon correlations. We show numerically that this process is well-defined quantum mechanically, and we explain the quantum phase space structures produced by such parametric amplification.

I. INTRODUCTION

In this paper, we discuss a generalization of squeezed states. Throughout this paper, "squeezing" refers to ordinary squeezing, and the term "generalized squeezing" shall be used explicitly when referring to our work.

In experiments where all sources of external noise have been made insignificant, there are still limits in measurement precision that are due to the Heisenberg uncertainty principle. This uncertainty in the variables is like a "quantum noise." Now we can use an idealized prescription in which a detector is coupled to a harmonic oscillator. This is quite a general prescription (e.g., the harmonic oscillator's coordinates could represent the electric field in light, or the position or momentum in a mass-spring system), so we shall restrict our attention to such systems. The quantum states that most closely describe the classical motion of these systems are harmonic oscillator coherent states. If we call the canonically conjugate variables for our harmonic oscillator system "position" x and "momentum" p, but choose appropriate dimensionless units for them, then the coherent states have a symmetric uncertainty in x and p, with $\Delta x \Delta p = 1$ and $\Delta x = \Delta p = 1$.

A loose classical description of quantum states can be made in terms of a phase-space probability distribution. For coherent states, this corresponds to a bivariate Gaussian distribution in x and p displaced from the origin and rotating around it with

time. The rotation in phase space corresponds to the oscillation of p one quarter cycle ahead of x, and the equality of the uncertainties in x and p initially leads to both Δx and Δp being independent of time. Thus, the harmonic oscillator coherent state has a time-stationary "noise."

Since the uncertainty principle puts a restriction only on the product $\Delta x \Delta p$, we might consider starting with a more precise position, so $\Delta x < 1$, and a more uncertain momentum, $\Delta p > 1$. These states are called squeezed states;[1] the noise has been squeezed into one variable at the expense of its conjugate. We see that as this state evolves freely in a harmonic oscillator potential, the precise x rotates in phase space to become a precise p in one-quarter cycle, and the uncertainty in p gets rotated into that of x. This variation of the uncertainty in these variables gives squeezed states a time-dependent noise (or as we shall refer to it in this paper, a phase-dependent noise).

A simple change in the variable we observe can yield a time-stationary noise for squeezed states. Roughly, this change corresponds to measuring either x or p in a rotating frame in phase space. This new variable is called the quadrature phase (or quadrature amplitude), so squeezed states have a time-stationary quadrature-phase noise. In practice, the quadrature phase is measured by interfering the original signal with an oscillating reference. This is just a homodyne or heterodyne detection scheme.

Real two-photon devices have been used recently to produce squeezed states of light.[2,3] With the use of these devices comes the promise of improved, high-precision interferometers.[4–6]

Any states deserving the name "generalized squeezed states" should have properties analogous to ordinary squeezed states; they must reduce to ordinary squeezing appropriately, and they should be generated from a phenomenologically reasonable

model. The first requirement is somewhat vague, so we shall simply take it to mean that these new states should at least possess a phase-dependent noise. We concentrate on a class of devices that generate states satisfying these criteria.

In Section II, such devices are modeled by ideal k-photon "parametric amplifiers" (creating or destroying k photons at a time) acting, for simplicity, on a single mode of the electromagnetic field. Ordinary squeezing corresponds to $k=2$. This section also analyzes these devices in a qualitative classical manner and suggests why we expect them to generate phase-dependent noise.

A simple but quantitative way to discuss the noise of these generalized squeezed states is through the use of a quantum analogue of the joint probability distribution in phase space. This analogue is known as the Q-function. We review this quantum description in Section III and derive some analytic properties that the Q-function of our k-photon state must satisfy. We find that the analytic properties of the Q-function closely follow our qualitative classical prejudices. In this section, we also look at the asymptotic behavior, for short time, of the Q-function. For the case $k=3$, we find that a large-amplitude coherent state is initially squeezed at a rate proportional to its amplitude—intrinsically much faster squeezing than that produced by ordinary squeezing interactions ($k=2$).

In Section IV, we study this classical-quantum correspondence more closely. We start by finding the proper classical Hamiltonian corresponding to these k-photon devices. Having set up the classical problem, we find that the classical and quantum evolutions differ for $k>1$. The classical motion is determined by an unstable fixed point at the origin of phase space; the classical evolution is "driven" by this fixed point to produce very sharp features in the classical probability distribution. These sharp features are destroyed by the quantum corrections, the dominant terms of which

correspond to ordinary diffusion when $k > 1$.

Fisher, Nieto, and Sandberg[7] have also studied this generalization of squeezing via k-photon devices. They concluded that there is something seriously wrong with the evolution operator for these devices, after discovering that for $k > 2$, its vacuum-to-vacuum matrix element has a divergent Taylor series expansion in time. In Section V, we study this matrix element numerically, using Padé approximants. We obtain good convergence for this matrix element for all scaled times (the coupling constant of the parametric amplifier multiplied by time) less than about 1.2. We also see that the divergence problems are due to singular behavior along the imaginary time axis. This matrix element and others are used to generate contour plots of the Q-function for $k = 3$ and $k = 4$ at various scaled times.

Hong and Mandel[8] have defined a set of parameters, which they call measures of "nth-order squeezing" for even n. When $n = 2$, this parameter reduces to a measure of the uncertainty in the quadrature phase and so corresponds to a measure of ordinary squeezing. We calculate the second-order and fourth-order squeezing parameters for some of the three-photon and four-photon states we have generated, and we find that these states are neither ordinarily squeezed nor squeezed to fourth order. Nonetheless, for the reasons already mentioned, we consider our states to be a generalization of squeezed states.

II. MOTIVATION FOR THE MODEL

One obvious generalization of squeezed states comes from recognizing that coherent states and squeezed states can be generated from idealized, one-photon and two-photon devices, respectively.

We restrict our attention to a single mode of the electromagnetic field, at frequency ω, which can be represented by the harmonic oscillator annihilation and creation operators (a and $a^\dagger$, respectively.) In the Schrödinger picture, the Hamiltonians for ideal one-photon and two-photon devices are given by

$$H_1 = \omega a^\dagger a + i[z_1(t)a^\dagger - z_1^*(t)a], \tag{2.1}$$

$$H_2 = \omega a^\dagger a + i[z_2(t)a^{\dagger 2} - z_2^*(t)a^2], \tag{2.2}$$

where

$$z_1(t) = |z| \exp[i(\phi - \omega t)], \tag{2.3}$$

$$z_2(t) = |z| \exp[i(\phi - 2\omega t)], \tag{2.4}$$

are time-dependent, complex, coupling constants. Throughout the following, unless we write $\hbar$ explicitly we shall use units for which $\hbar = 1$. We have chosen the time dependence of the couplings z_1 and z_2 so that, in the interaction picture, the Hamiltonians are independent of time; this corresponds to running the devices at resonance.

In the interaction picture, the time-evolution operators are

$$U_1 = \exp[(za^\dagger - z^* a)t], \tag{2.5}$$

$$U_2 = \exp[(za^{\dagger 2} - z^* a^2)t], \tag{2.6}$$

where $z = |z| \exp(i\phi)$ is a time-independent coupling constant. The operators U_1 and U_2 correspond to the displacement operator and the squeezing operator, respectively. U_1 and U_2 are already time-ordered here, since H_1 and H_2 are time-independent in this picture.

In this paper we study an idealized, degenerate, k-photon parametric amplifier, which reduces to Eqs. (2.1) and (2.2) for $k=1$ and $k=2$, respectively. In the Schrödinger picture, its Hamiltonian is

$$H_k = \omega a^\dagger a + i[z_k(t)a^{\dagger k} - z_k^*(t)a^k]$$

$$\sim \omega a^\dagger a + 2|z|\sin(k\omega t - \phi)E^k \ , \tag{2.7}$$

where $z_k(t) = |z|\exp[i(\phi - k\omega t)]$. The second form in Eq. (2.7) requires the rotating-wave approximation, and we have taken the electric field operator to be

$$E \sim (a^\dagger + a) \ . \tag{2.8}$$

For the rest of the paper, we shall work in the interaction picture, so the time evolution operator corresponding to Eq. (2.7) is

$$U_k(z, z^*; t) = \exp[(za^{\dagger k} - z^* a^k)t] \ , \tag{2.9}$$

where $z \equiv |z|\exp(i\phi)$.

The second form in Eq. (2.7) allows the classical interpretation for the interaction as a nonlinear force oscillating at k times the natural frequency of the optical mode to which it is coupled. (For $k=1$, this is resonant excitation. For $k=2$, this is like kicking on a swing, where we force at twice the fundamental frequency.) This will excite the mode where its phase differs by $0, \cdots, 2\pi(k-1)/k$ radians from the phase of the "force" $z_k(t)$ and will damp it at the intermediate phases. Thus, for $k > 1$, we may expect this interaction to produce phase-dependent noise.

Equation (2.7) also suggests a realization of the interaction, for which we make use of the macroscopic description for the electromagnetic field with matter; this is given by the polarizability

$$P_i = \chi_{ij}^{(1)} E_j + \chi_{ijm}^{(2)} E_j E_m + \cdots \,. \tag{2.10}$$

Here, $\chi_{ij}^{(1)}$ is the first-order susceptibility of the material (which is simply $\chi^{(1)}\delta_{ij}$ for an isotropic material), $\chi^{(n)}$ is the nth-order nonlinear susceptibility, and the subscripts correspond to spatial and optical polarization components. The electromagnetic energy density is

$$\frac{1}{8\pi}\sum_i [(E_i + 4\pi P_i)E_i + \text{magnetic terms}] \,, \tag{2.11}$$

which from the decomposition in Eq. (2.8) has terms

$$\omega a^\dagger a + \cdots + \frac{1}{2}\chi^{(k)}(a^{\dagger k} + a^k)E_{\text{ext}}(t) + \cdots \,, \tag{2.12}$$

where E_{ext} is an external "pump" field, and we have ignored all but one spatial component and one optical polarization. A comparison with Eq. (2.7) yields

$$z(t) \sim \frac{-i}{2}\chi^{(k)}E_{\text{ext}}(t)V \,, \tag{2.13}$$

where V is the volume factor corresponding to the volume of the nonlinear material. (Similar terms arise from higher-order susceptibilities.)

This discussion shows how the model interaction in Eq. (2.7) may be formed by a kth-order (or higher) susceptibility. Of course, as Eq. (2.7) is only a model for such interactions, some details are left out. We have assumed an ideal pump for which there is no depletion. We neglect depletion, since we expect only significant nonlinearities for very intense sources. We are also neglecting the dissipation and fluctuations that go with the mode being coupled to a thermal bath. Finally, we are glossing over the microscopic nature of the nonlinearity and quantizing directly the

macroscopic equations of classical electromagnetism.

III. ANALYTIC PROPERTIES

Throughout this paper, we are interested in the properties of the states generated from the interactions of Eq. (2.7). To describe these properties, we shall make use of a quantum description that is closely linked to the phase space distribution for a classical system.

The description we use, called the Q-function, is itself a probability distribution whose moments give "antinormally" ordered expectation values[9]

$$<a^n a^{\dagger m}> \equiv \mathrm{tr}[a^n a^{\dagger m} \rho(t)] = \int \frac{d^2\alpha}{\pi} \alpha^n \alpha^{*m} Q(\alpha, \alpha^*; t) , \tag{3.1}$$

where $d^2\alpha \equiv d\,\mathrm{Re}(\alpha)\, d\,\mathrm{Im}(\alpha)$, and

$$\rho(t) = U_k(t) \rho(0) U_k^{\dagger}(t) . \tag{3.2}$$

Here, $U_k(t)$ is the time-evolution operator [Eq. (2.9)] for our k-photon device, and $\rho(t)$ is the density operator for the state that determines the Q-function via[9,10]

$$Q(\alpha, \alpha^*; t) \equiv <\alpha | \rho(t) | \alpha> = \mathrm{tr}[\rho(t) \delta(a - \alpha) \delta(a^{\dagger} - \alpha^*)] , \tag{3.3}$$

where $|\alpha>$ is the harmonic oscillator coherent state, $a\,|\alpha> = \alpha\,|\alpha>$. Equation (3.3) may be inverted to give

$$\rho(t) = \int \frac{d^2\alpha}{\pi} \delta(a^{\dagger} - \alpha^*) \delta(a - \alpha) Q(\alpha, \alpha^*; t) . \tag{3.4}$$

The expressions with operator-valued Dirac-delta functions may be considered as purely formal expressions for repeated Fourier transformations; we may write them as

$$\delta(a^\dagger-\alpha^*)\delta(a-\alpha)\equiv\int\frac{d^2\xi}{\pi}\exp[i\xi(a^\dagger-\alpha^*)]\exp[i\xi^*(a-\alpha)]\,,$$

$$\delta(a-\alpha)\delta(a^\dagger-\alpha^*)\equiv\int\frac{d^2\xi}{\pi}\exp[i\xi^*(a-\alpha)]\exp[i\xi(a^\dagger-\alpha^*)]\,. \tag{3.5}$$

The Q-function also represents[11] the joint probability distribution for a specially set-up and balanced detector making simultaneous measurements of ''position'' $a^\dagger+a$ and ''momentum'' $i(a^\dagger-a)$. Hence, we shall refer to contour plots of Q on $\mathrm{Im}(\alpha)$ versus $\mathrm{Re}(\alpha)$ as ''phase-space'' diagrams.

Let us define the rotation operator

$$R(\Theta)\equiv\exp(-i\Theta a^\dagger a)\,; \tag{3.6}$$

then

$$R(\Theta)|\alpha\rangle=|\alpha e^{-i\Theta}\rangle\,, \tag{3.7}$$

and

$$a_\Theta\equiv R^\dagger(\Theta)aR(\Theta)=ae^{-i\Theta}\,. \tag{3.8}$$

Applying this rotation to $\rho(t)$ is equivalent to taking $\alpha\rightarrow\alpha\exp(-i\Theta)$ in $Q(\alpha,\alpha^*;t)$, which corresponds to rotation of the phase-space diagram. Also, $R(\Theta)$ transforms the time-evolution operator $U_k(t)$ to

$$R^\dagger(\Theta)U_k(t)R(\Theta)=\exp[(ze^{ik\Theta}a^{\dagger k}-z^*e^{-ik\Theta}a^k)t]\,; \tag{3.9}$$

i.e., $z\rightarrow z\exp(ik\Theta)$. When the initial state is the vacuum, $\rho(0)=|0\rangle\langle 0|$, a rotation by $\Theta=2\pi/k$ maps Q to itself, so that the phase-space plot will have a k-fold symmetry. This is familiar for $k=1$ and $k=2$, the coherent and squeezing cases.

For convenience, we shall henceforth restrict the coupling constant z to be real; when the initial state is the vacuum, this corresponds simply to a rotation of the phase-space diagram. We may now define a scaled time, $r \equiv |z|t$, so that the time-evolution operator reduces to

$$U_k(r) = \exp[r(a^{\dagger k} - a^k)] . \tag{3.10}$$

We now derive the evolution equation for the Q-function. Although this is standard,[10] we present it for completeness. We start with the equation of motion for the density operator in the interaction picture:

$$i\frac{d\rho}{dt} = i|z|[a^{\dagger k} - a^k, \rho] . \tag{3.11}$$

Using the last expression in Eq. (3.3) and the relations

$$[a, \delta(a^\dagger - \alpha^*)] = \left\{ \frac{\partial}{\partial a^\dagger}\delta(a^\dagger - \alpha^*) \right\} = -\frac{\partial}{\partial \alpha^*}\delta(a^\dagger - \alpha^*) , \tag{3.12}$$

$$[a^\dagger, \delta(a - \alpha)] = \left\{ -\frac{\partial}{\partial a}\delta(a - \alpha) \right\} = \frac{\partial}{\partial \alpha}\delta(a - \alpha) , \tag{3.13}$$

$$a\,\delta(a - \alpha) = \alpha\delta(a - \alpha) , \tag{3.14}$$

$$a^\dagger \delta(a^\dagger - \alpha^*) = \alpha^*\delta(a^\dagger - \alpha^*) , \tag{3.15}$$

we find that the Q-function evolves according to

$$\frac{\partial Q}{\partial r} = LQ \equiv \left[\alpha^k - \left(\alpha + \frac{\partial}{\partial \alpha^*} \right)^k + \alpha^{*k} - \left(\alpha^* + \frac{\partial}{\partial \alpha} \right)^k \right] Q , \tag{3.16}$$

where L is the Liouvillian. Equation (3.16) has the formal solution

$$Q(r)=e^{rL}Q(0)=Q(0)+LQ(0)r+\frac{1}{2}L^2Q(0)r^2+\cdots, \tag{3.17}$$

where $Q(0)$ is the initial value of the Q-function.

If we treat this perturbation series for $Q(r)$ as an asymptotic series (Sec. V), then it is indeed valid to truncate the series to find the asymptotic behavior for small time. For the initial state in the vacuum $\rho(0)=|0><0|$, corresponding to $Q(0)=\exp(-|\alpha|^2)$, we find

$$\begin{aligned} Q(r) &\sim e^{-|\alpha|^2}[1+r(\alpha^k+\alpha^{*k})]+O(r^2) \\ &= e^{-|\alpha|^2}[1+2r|\alpha|^k\cos(k\theta)]+O(r^2), \end{aligned} \tag{3.18}$$

where $\alpha=|\alpha|e^{-i\theta}$. Thus, for $r \ll 1$, the Q-function develops k lobes along $\theta=0, 2\pi/k, \cdots, 2\pi(k-1)/k$ and dips between these lobes. We shall see later (Section V) that this is in agreement with more detailed calculations and even for such short times is different from the classical behavior (Section IV).

Similarly, if we start in an initial coherent state with real amplitude x_0, so that $\rho(0)=|x_0><x_0|$ and $Q(0)=\exp(-|\alpha-x_0|^2)$, then to first order in r,

$$Q(r)\sim e^{-|\alpha-x_0|^2}[1+r(\alpha^k+\alpha^{*k}-2x_0^k)]+O(r^2),$$

and for $x_0 \gg |\alpha-x_0|$,

$$Q(r)\sim\exp\left[-\frac{[\mathrm{Re}(\alpha)-x_0-rkx_0^{k-1}]^2}{1+rk(k-1)x_0^{k-2}}-\frac{[\mathrm{Im}(\alpha)]^2}{1-rk(k-1)x_0^{k-2}}\right]+O(r^2). \tag{3.19}$$

Thus, $\mathrm{Re}(\alpha)$ and $\mathrm{Im}(\alpha)$ have mean values and uncertainties given by

$$\sigma_{\mathrm{Re}(\alpha)} \sim \frac{1}{\sqrt{2}}\left[1+\frac{r}{2}k(k-1)x_0^{k-2}\right]+O(r^2), \tag{3.20}$$

$$\sigma_{\mathrm{Im}(\alpha)} \sim \frac{1}{\sqrt{2}}\left[1-\frac{r}{2}k(k-1)x_0^{k-2}\right]+O(r^2), \tag{3.21}$$

$$<\mathrm{Re}(\alpha)> \sim x_0+rkx_0^{k-1}+O(r^2), \tag{3.22}$$

$$<\mathrm{Im}(\alpha)> \equiv 0, \tag{3.23}$$

for $r \ll 1$; the standard deviations $\sigma_{\mathrm{Re}(\alpha)}$ and $\sigma_{\mathrm{Im}(\alpha)}$ are independent of x_0 only for $k=1$ and $k=2$. For $k=3$, a resonant 3-photon parametric amplifier will initially squeeze an in-phase coherent state at a rate proportional to its amplitude. Also, there would be no ordinary squeezing of the vacuum for any of the $k>2$ processes. Similar results have been obtained by Hillery, Zubairy and Wódkiewicz.[12] They have also shown that this generalized parametric amplifier produces no ordinary squeezing of the vacuum when $k > 2$, even for large r.

IV. CLASSICAL DYNAMICS

It is worthwhile calculating the classical dynamics associated with the interaction in Eq. (2.7) in detail, not only to confirm our intuition from Section II, but also to compare it to the full quantum dynamics.

There is not always, however, a unique classical system that one might associate with a quantum Hamiltonian. In order to choose a classical Hamiltonian, we start by writing the quantum theory in terms of a path integral; the classical Hamiltonian then appears as part of the classical action in the path integral. Another route to the classical Hamiltonian may be found in Milburn.[13]

In the coherent state representation of Klauder,[14] the full propagator is

$$<\alpha_f;t\,|\,\alpha_i;0>=\int_{\alpha(0)=\alpha_i}^{\alpha(t)=\alpha_f} D(\alpha(\tau))\exp\left[\int_0^t d\tau[\frac{1}{2}(\alpha\dot{\alpha}^*-\alpha^*\dot{\alpha})-iH_N(\alpha^*,\alpha;\tau)]\right], \quad (4.1)$$

where $D(\alpha(\tau))$ is the path integral measure, and

$$H_N(\alpha^*,\alpha;t)\equiv<\alpha|H(t)|\alpha>=\omega|\alpha|^2+i|z|(\alpha^{*k}e^{-ik\omega t}-\alpha^k e^{ik\omega t}) \quad (4.2)$$

is the appropriate classical Hamiltonian, which is just the quantum Hamiltonian in normal order form, with a replaced by α and $a^\dagger$ by α^*. The Poisson bracket then yields the classical equation of motion

$$\frac{\partial Q_{\rm cl}}{\partial t}=-\{H_N,Q_{\rm cl}\}=i\frac{\partial H_N}{\partial\alpha^*}\frac{\partial Q_{\rm cl}}{\partial\alpha}-i\frac{\partial H_N}{\partial\alpha}\frac{\partial Q_{\rm cl}}{\partial\alpha^*}, \quad (4.3)$$

where $Q_{\rm cl}$ is the classical probability distribution in phase space.

In a rotating frame (analogous to the interaction picture), the classical equation (after scaling out the coupling constant $|z|$) is therefore

$$\frac{\partial Q_{\rm cl}}{\partial r}=-k\left[\alpha^{k-1}\frac{\partial Q_{\rm cl}}{\partial\alpha^*}+\alpha^{*k-1}\frac{\partial Q_{\rm cl}}{\partial\alpha}\right]. \quad (4.4)$$

It has the asymptotic dynamics

$$Q_{\rm cl}=e^{-|\alpha|^2}[1+2kr|\alpha|^k\cos(k\theta)]+O(r^2), \quad (4.5)$$

for an initial $Q_{\rm cl}$ that corresponds to Q for the vacuum; i.e., $Q_{\rm cl}=\exp(-|\alpha|^2)$. The lobes of $Q_{\rm cl}$ grow k times faster than the corresponding quantum features [see Eq. (3.18)], at least for $r \ll 1$.

We can make an interesting general observation about these classical equations of motion at this point. Writing the quantum equation of motion in terms of $H_N(\alpha^*,\alpha)$, we have

$$i\frac{\partial Q}{\partial t} = H_N\left(\alpha^*,\alpha+\frac{\vec{\partial}}{\partial\alpha^*}\right)Q - QH_N\left(\alpha^*+\frac{\tilde{\partial}}{\partial\alpha},\alpha\right) \tag{4.6}$$

$$= \frac{\partial H_N}{\partial\alpha}\frac{\partial Q}{\partial\alpha^*} - \frac{\partial H_N}{\partial\alpha^*}\frac{\partial Q}{\partial\alpha} + (\text{higher derivatives of } Q)\,, \tag{4.7}$$

where the derivative $\tilde{\partial}/\partial\alpha$ acts to the left. Thus, truncating the full quantum description of Eq. (4.6) to first order in the derivatives of Q leads us to the same classical equation we derived from the path integral formulation [Eq. (4.3)]. This truncation obviously leads to classical dynamics, and the classical trajectories are given by solving the characteristic equations for the resulting first-order partial differential equation. These characteristic equations (classical trajectories) for Eq. (3.16) are

$$\frac{d\alpha}{dr} = k\alpha^{*k-1}\,, \tag{4.8}$$

$$\frac{d\alpha^*}{dr} = k\alpha^{k-1}\,. \tag{4.9}$$

These classical trajectories conserve

$$H_N \propto \frac{i}{2}(\alpha^{*k}-\alpha^k)\,, \tag{4.10}$$

so the classical trajectories are curves of constant $\mathrm{Im}(\alpha^k)$. These curves describe a flow in phase space around an unstable fixed point, with k directions of damping and k directions of growth. Thus, the Hamiltonian has a k-saddle around the unstable

fixed point. This is a generalization of the generic saddle of unstable fixed points.

Integrating the characteristic equations numerically, for $k=3$, gives Figs. 1(a) and 1(b); the threefold symmetry and the development of lobe structure are apparent. We also see that arbitrarily sharp structure appears as r increases. This sharp structure is a feature of the classical dynamics of many systems.[15] We do not expect to see it in the quantum system, for as it develops, the higher derivatives become significant. For instance, the sharp structure develops along the lines $\theta=0, 2\pi/k, \cdots, 2\pi(k-1)/k$; so expanding Eq. (4.4) about $\theta=0$ with

$$|\alpha| \gg \max\left[1, \frac{1}{|\alpha|}\left|\frac{\partial Q}{\partial \theta}\right|\right],$$

and

$$\frac{1}{|\alpha|}\left|\frac{\partial Q}{\partial \theta}\right| \gg \left|\frac{\partial Q}{\partial |\alpha|}\right|, \tag{4.11}$$

one finds that for any k the term that dominates is the diffusion term, leading to the equation

$$\frac{\partial Q}{\partial r} = \frac{1}{4}k(k-1)|\alpha|^k - 2\left[\frac{1}{|\alpha|^2}\frac{\partial^2 Q}{\partial \theta^2}\right]. \tag{4.12}$$

We see that this diffusion term vanishes for $k=1$ and is independent of $|\alpha|$ for $k=2$. For $k>2$, it is proportional to $|\alpha|^k - 2$; thus any attempt to squeeze at a faster rate as suggested at the end of Section III will have to fight this diffusion. Numerical results suggest that even for $k=4$, this diffusion dominates.

It is worth noting that this "quantum diffusion" is not the only way quantum mechanics may prevent fine structures being formed by the classical evolution.

Milburn[13] has studied a different nonlinear potential, which produces a negative-definite diffusion term; it prevents the classical structure by forcing recurrences of the initial state.

V. QUANTUM DYNAMICS

There are at present no analytic techniques for dealing with the full problem. One obvious method would be to normal-order the evolution operator U_k [Eq. (3.10)]; however, for $k > 2$, this is an unsolved problem.[16] Fisher, Nieto, and Sandberg[7] look at the matrix element

$$<0|U_k(r)|0> = 1 - \frac{k!}{2!}r^2 + \frac{(k!)^2 + (2k)!}{4!}r^4$$

$$- \frac{(k!)^3 + 2k!(2k)! + (2k)!^2/k! + (3k)!}{6!}r^6$$

$$+ \cdots + (-1)^n \frac{(nk)!}{(2n)!}C_n r^{2n} + \cdots \tag{5.1}$$

by expanding it as a Taylor series in r. All the coefficients C_n are positive. They point out that this series has a zero radius of convergence; in fact, the coefficients C_n grow like β^n, where $\beta(k)$ is near 1. Thus, the vacuum state is by definition[17] a non-analytic vector with respect to the Hamiltonian H_k. This does not, however, reflect on the existence of the Hamiltonian, as they imply, or on the unitarity of U_k. The non-convergence is, in fact, a mathematical artifact caused by singular behavior on the imaginary time axis, since the Taylor series converge only up to the nearest pole. (This phenomenon is quite common in parabolic equations; for example, the initial value problem for the heat equation has no solution for negative time.)

There are many methods of analytic continuation available, which obtain useful information from the Taylor series outside their radius of convergence. If that radius is actually zero, then we can still hope to extract information by deriving an alternative form that does converge, since in this case the Taylor expansion would be an asymptotic series about $r = 0$. One such method is that of Padé approximants,[18,19] in which the $[N/M]$ approximant is a rational function $P_M^N(r^2) = P_N(r^2)/Q_M(r^2)$, where P_N (Q_M) is a polynomial of degree N (M). The coefficients of the polynomials P_N and Q_M are chosen so as to match the first $M + N + 1$ Taylor coefficients of the function. (Q_M has leading coefficient 1.) Padé approximants analytically continue well because they tend to reproduce the pole structure that limits the Taylor series' convergence. A common application is to form a diagonal series from the $[N/N]$ and $[N/N+1]$ approximants (these require $2N + 1$ and $2N + 2$ coefficients, respectively); this often has remarkable convergence properties. We restrict our attention to this series.

In practice, one represents the rational function as a continued fraction. Surprisingly, this means that successive Padé approximants can be calculated by generating only one additional continued-fraction coefficient.[18] This breaks down if there is a zero Taylor coefficient. Hence, we regard Eq. (5.1) as an expansion in r^2; i.e., we analytically continue in r^2 instead of r. The algorithms tend to be numerically sensitive—about half a digit is lost for every extra term required—so we worked in high precision throughout (~ 33 significant figures, with checks at even higher accuracy). This is not a problem with Padé approximants, but rather a characteristic of the numerical treatment of analytic continuation methods in general, which try to predict behavior far from the origin from the first few terms of the Taylor series. For instance, in this case to perform Borel summability instead of Padé approximation requires the

numerical integration of rapidly oscillating functions.

For $k=1$ and $k=2$, the series of Padé approximants converges rapidly to the known solutions

$$<0|U_1(r)|0> = \exp(-r^2/2)\,, \tag{5.2}$$

$$<0|U_2(r)|0> = \cosh^{-1/2}(2r)\,. \tag{5.3}$$

For $k=3$, we see convergence out to $r \sim 1.2$ [see Fig. 2(a)]. One reason for this is that the first 56 continued fraction coefficients are all positive, which gives the relations[18]

$$P^N_{N+1} \leq P^{N+1}_{N+2} \leq P^{N+1}_{N+1} \leq P^N_N \tag{5.4}$$

for $2N+2 \leq 56$; i.e., the odd and even Padé approximants bracket their limit [see Fig. 2(b)].

P^N_M, in general, has M simple poles, some of which mimic the real poles of the function and some of which are spurious. The convergence of the approximants will then depend on whether the spurious poles are eventually canceled by zeros in the numerator or move off to infinity as N and $M \rightarrow \infty$. In our case, the poles initially lie on the negative real axis (corresponding to imaginary r), clustering close to zero; this is the usual behavior on a branch cut. However, coefficients 57, 58, and 168 are negative, introducing spurious poles. These prevent convergence for large r, unless enough terms are included to move the pole well past r [see Figs. 2(b) and 3].

We repeated these calculations for different matrix elements, enabling us to compute the Q-function. For a system initially in the vacuum state, $\rho(0) = |0><0|$, a decomposition in the number state basis yields

$$Q(r) = \exp(-|\alpha|^2)\left|\sum_{n=0}^{\infty} \frac{\alpha^{*n}}{\sqrt{n!}} <n|U_k(r)|0>\right|^2 . \tag{5.5}$$

We note the the finiteness of the vacuum-to-vacuum matrix element of Eq. (5.1) determines the finiteness of all the other matrix elements in the sum of Eq. (5.5). This is because these matrix elements are nonzero only when n is a multiple of k, in which case they can be written as sums of derivatives of the vacuum-to-vacuum matrix element:

$$<k\,|\,U_k(r)\,|\,0> = \frac{1}{\sqrt{k!}}\frac{d}{dr}<0\,|\,U_k(r)\,|\,0>\,,$$

$$<2k\,|\,U_k(r)\,|\,0> = \frac{1}{\sqrt{(2k)!}}\left[k! + \frac{d^2}{dr^2}\right]<0\,|\,U_k(r)\,|\,0>\,, \tag{5.6}$$

etc. Now $|<n\,|\,U_k(r)\,|\,0>|^2$ is a probability, so it is bounded above by 1. Thus, the sum in Eq. (5.5) converges rapidly; e.g., for $\alpha \lesssim 5$, we need to include only terms up to $n = 74$ (this corresponds to only 25 nonzero terms for $k = 3$) in the sum for each additional term, to give a contribution of less than one percent. The contour plots of the Q-function are shown in Figs. 1(c)–1(e) for $r = 0.05$, 0.2, and 1.0. Notice the threefold symmetry and the lobe development, which is much slower and wider than in the classical case (Sections III, IV). Our techniques also work for higher k [see Fig. 1(f) for $k = 4$], but the numerical problems become more extreme.

As a final point, we calculated two of Hong and Mandel's[8] "higher-order squeezing" parameters, $<:(\Delta a_{\theta 1})^{2n}:>$, for the quantum states in Figs. 1(c)–1(f). Here, the colons denote normal ordering, and $a_{\theta 1} = a_\theta + a_\theta^\dagger$ is the quadrature amplitude defined at an angle θ to the Re(α) axis [Eq. (3.8)]. Using the P representation[9,10]

$$P(\alpha, \alpha^*) = \mathrm{tr}[\rho(t)\delta(a^\dagger - \alpha^*)\delta(a - \alpha)]\,, \tag{5.7}$$

which generates the expectation values of normally ordered operators, and the relation

for reordering the Dirac-delta functions

$$\delta(a^\dagger-\alpha^*)\delta(a-\alpha)=\exp\left[-\frac{\partial^2}{\partial\alpha\partial\alpha^*}\right]\delta(a-\alpha)\delta(a^\dagger-\alpha^*)\,, \tag{5.8}$$

we find

$$P(\alpha,\alpha^*)=\exp\left[-\frac{\partial^2}{\partial\alpha\partial\alpha^*}\right]Q(\alpha,\alpha^*)\,. \tag{5.9}$$

Hence, the normally ordered ''squeezing'' parameter becomes

$$\begin{aligned}<:(\Delta a_{\Theta 1})^{2n}:> &= \int\frac{d^2\alpha}{\pi}P(\alpha,\alpha^*)(\alpha_\Theta+\alpha_\Theta^*-<a_{\Theta 1}>)^{2n} \\ &= \int\frac{d^2\alpha}{\pi}Q(\alpha,\alpha^*)\exp\left[-\frac{\partial}{\partial\alpha\partial\alpha^*}\right](\alpha_\Theta+\alpha_\Theta^*-<a_{\Theta 1}>)^{2n}\,,\end{aligned} \tag{5.10}$$

after integration by parts.

These parameters are trivially positive for classical states, and Hong and Mandel call a state squeezed to $2n$th order if $<:(\Delta a_{\Theta 1})^{2n}:>$ is negative. In each case described in Figs. 1(c)–1(f), we calculated this parameter for $n=1$ and $n=2$, with $\Theta=0,\pi/k$, and found it to be positive in each case; i.e., there is no fourth-order squeezing of the kind described by Hong and Mandel[8], nor is there any ordinary squeezing for these states (ordinary squeezing occurs when their parameter for $n=1$ is negative). One reason why we might have expected this is that their parameters make use of an orthogonal decomposition of phase space into quadrature amplitudes; instead, our states have a k-fold symmetry, which will not match this decomposition when $k>2$.

As pointed out by Fisher, Nieto, and Sandberg,[7] the self-adjointness and Hermiticity of the interaction Hamiltonian $H_k = i(za^{\dagger k} - z^* a^k)$ are not in question. What is questioned is the unitarity of the theory, but the only potential problem we see would be if the matrix elements of the time evolution operator $U_k(r)$ did not exist. We have demonstrated numerically that they do.

VI. CONCLUSIONS

We conclude that this generalization of ''one-photon'' coherent and ''two-photon'' squeezed states to ''many-photon'' states is possible and that these non-Gaussian states show quantum features quite different from the classical approximation. These k-photon interactions for $k > 2$ initially generate ordinary squeezing at a higher rate than the usual $k = 2$ parametric amplifier when they act resonantly on a large amplitude coherent state. This occurs, however, in competition with a quantum diffusion that gets stronger for successively larger k.

ACKNOWLEDGMENTS

One of the authors (S.L.B.) would like to thank Carlton M. Caves for pointing out Ref. 7 and for many useful discussions on the physics; he also appreciates discussions with Daniel Neuhauser on high-precision calculations. R.I.M. would like to thank Dan Meiron for discussions on analytic continuation. This work was supported in part by the Office of Naval Research [Contract No. N00014-85-K-0005]. R.I.M. was supported in part by the University Grants Committee of New Zealand.

REFERENCES

[1] For a review, see D. F. Walls, Nature **306**, 141 (1983).

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

[3] See A. L. Robinson, Science **233**, 280 (1986).

[4] C. M. Caves, Phys. Rev. D **23**, 1693 (1981).

[5] R. S. Bondurant and J. H. Shapiro, Phys. Rev. D **30**, 2548 (1984).

[6] B. Yurke, J. R. Klauder and S. L. McCall, Phys. Rev. D **33**, 4033 (1986).

[7] R. A. Fisher, M. M. Nieto and V. D. Sandberg, Phys. Rev. D **29**, 1107 (1984).

[8] C. K. Hong and L. Mandel, Phys. Rev. Lett. **54**, 323 (1985)

[9] K. E. Cahill and R. J. Glauber, Phys. Rev. **177**, 1882 (1969).

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

[11] E. Arthurs and J. L. Kelly, Jr., Bull. Syst. Tech. J. **44**, 725 (1965).

[12] M. Hillery, M. S. Zubairy and K. Wódkiewicz, Phys. Lett. **103*A***, 259 (1984).

[13] (a) G. J. Milburn, Phys. Rev. A **33**, 674 (1986); see also (b) G. J. Milburn and C. A. Holms, Phys. Rev. Lett. **56**, 2237 (1986).

[14] J. R. Klauder, Ann. Phys. (N.Y.) **11**, 123 (1960).

[15] M. V. Berry, in *Chaotic Behavior of Deterministic Systems, Les Houches Lectures, 1981*, edited by G. Iooss, R. H. G. Helleman, and R. Stora (North-Holland, Amsterdam, 1982).

[16] W. Witschell, Phys. Lett. **111*A***, 383 (1985).

[17] E. Nelson, Ann. Math. **70**, 572 (1959).

[18] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1978), p.383.

[19] G. A. Baker, *Essentials of Padé Approximants* (Academic, New York, 1975).

FIGURE CAPTIONS

FIG. 1. (a) and (b) Contour plots of the classical phase space distribution, $Q_{cl}(r)$, for $k=3$, at times (a) $r=0.05$ and (b) $r=0.2$. (c)–(f) Contour plot of the quantum Q-function, $Q(r)$, for (c) $r=0.05$ and $k=3$; (d) $r=0.2$ and $k=3$; (e) $r=1.0$ and $k=3$; and (f) $r=0.1$ for $k=4$.

FIG. 2. (a) Plot of the matrix element, $<0|U_{k=3}(r)|0>$, versus scaled time r, when calculated with $2N+1=167$ and $2N+2=168$ Padé coefficients. This shows convergence of the matrix element out to $r\sim 1.2$. (b) Plot of the value of the matrix element, $<0|U_{k=3}(r)|0>$, versus N, for various times: $r=0.05$; $r=0.2$; $r=0.4$; $r=0.7$; $r=1.0$; $r=1.2$; and $r=1.5$. P_N^N and P_N+1^N are plotted separately to show how they bracket their "limit" until they reach the spurious pole at $2N+2=58$. We also see that the spurious pole moves out to higher times as N is increased.

FIG. 3. Schematic plot of the positions of the poles in the Padé approximant to $<0|U_{k=3}(r)|0>$ as a function of r^2, for (a) $2N+2=40$; (b) $2N+2=60$; (c) $2N+2=120$; and $2N+2=160$. This shows the accumulation of poles near the origin (for $r^2<0$), which causes the difficultly with convergence of the Taylor series. It also shows the motion of the spurious pole, as in Fig. 2(b).

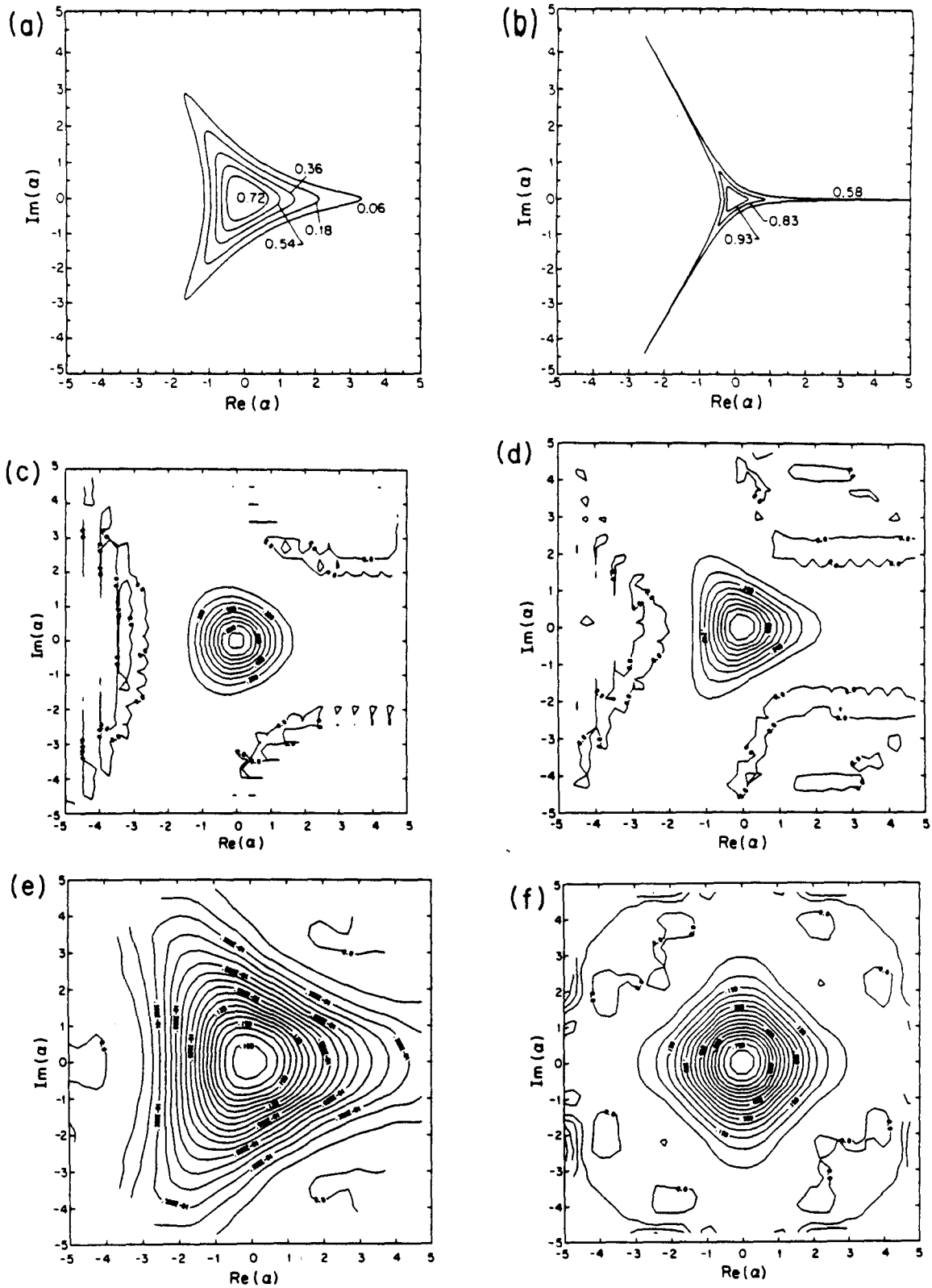
(a)
0.36
0.72
0.06
0.18
0.54
Im(α)
Re(α)
(b)
0.58
0.83
0.93
Im(α)
Re(α)
(c)
Im(α)
Re(α)
(d)
Im(α)
Re(α)
(e)
Im(α)
Re(α)
(f)
Im(α)
Re(α)

Figure 1

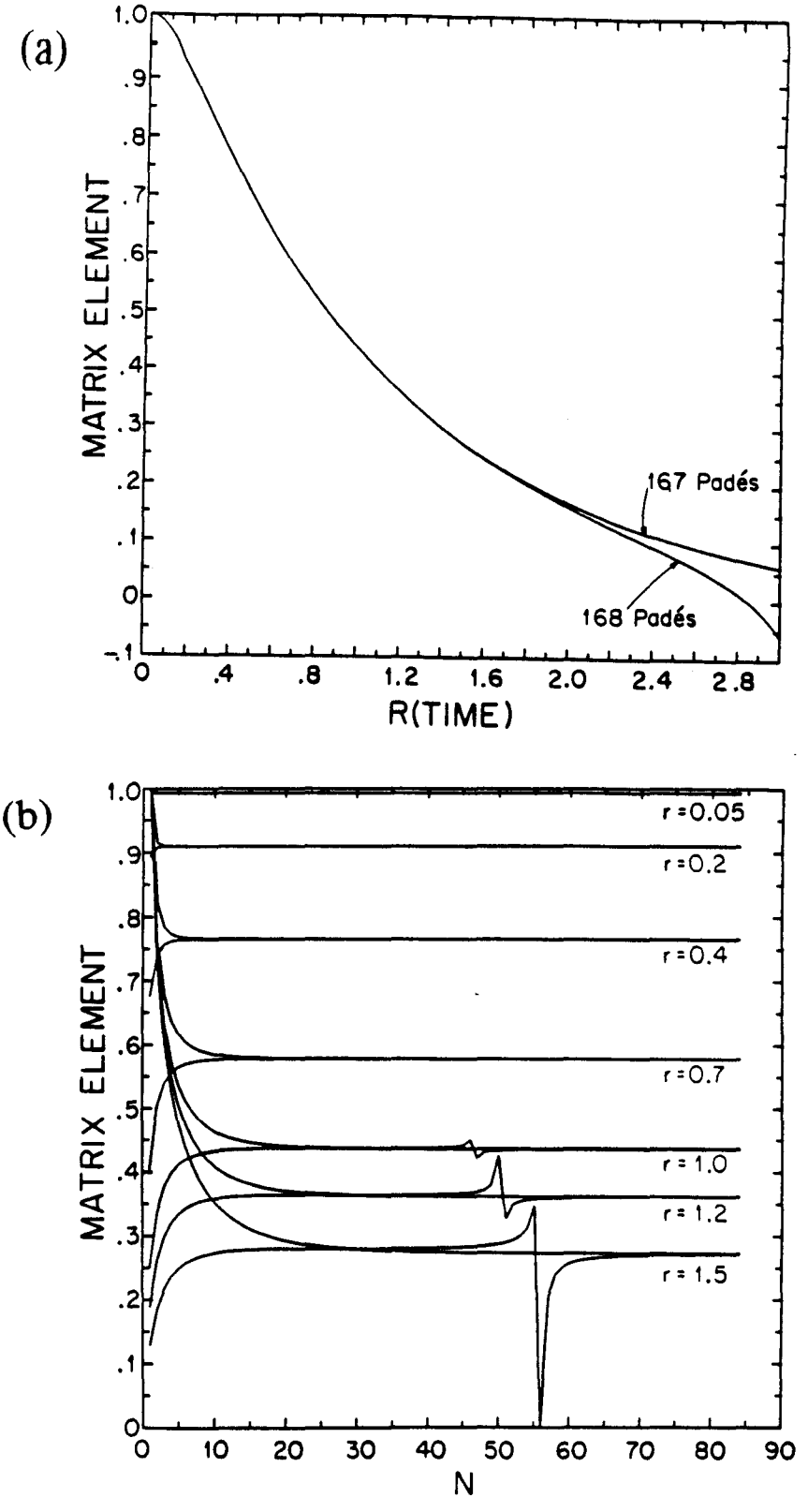
(a)
MATRIX ELEMENT
R(TIME)
167 Padés
168 Padés
(b)
MATRIX ELEMENT
N
r=0.05
r=0.2
r=0.4
r=0.7
r=1.0
r=1.2
r=1.5

Figure 2

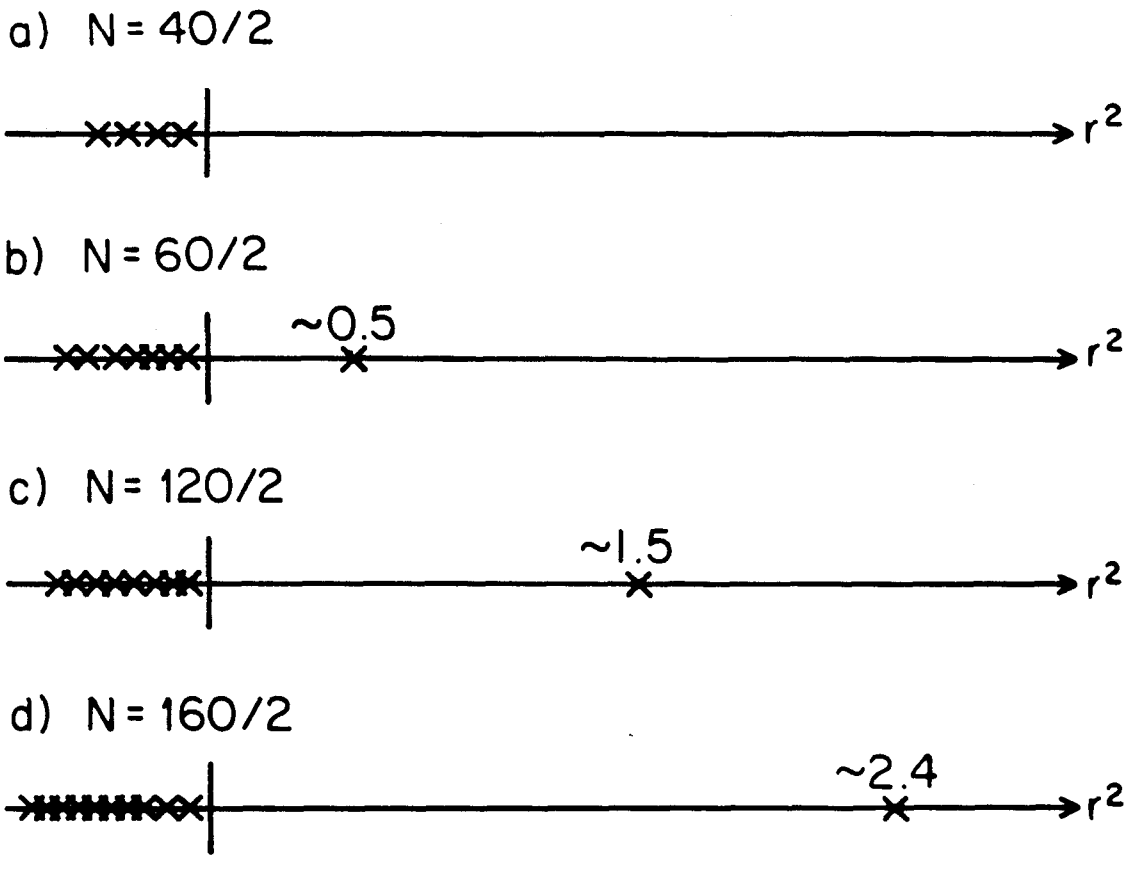

Figure 3

CHAPTER 3

Limitations to squeezing in a parametric amplifier due to quantum pump fluctuations

by Samuel L. Braunstein and David D. Crouch

Submitted to Physical Review A.

ABSTRACT

We perform discrete mode calculations for a parametric amplifier with a quantum pump, and discuss some of the limitations on calculations of this sort in quantum optics. We calculate corrections to the squeezing that is due to pump quantum fluctuations to order $1/\bar{N}^2$, for a pump initially in a coherent state with average photon number $\bar{N}$. We find that the limit to the variance of the squeezed quadrature that is due to the quantum nature of the pump goes as $\bar{N}^{-1/2}$.

I. INTRODUCTION

The parametric amplifier[1,2] (PA) is a basic device in quantum optics and quantum electronics. It couples a pump field at frequency ω_p to signal modes at frequencies near $\omega = \omega_p/2$. In this paper, we are mainly interested in the application of the PA for generating squeezed states,[3–5] *i.e.* quantum states for which one of a pair of canonically conjugate variables has its quantum noise (uncertainty) reduced below the vacuum level (zero point noise). The main purpose of this paper is to show that the ability of a PA to produce squeezed light is limited by the initial phase noise in the pump.

When the signal modes are initially in vacuum states, only the pump's phase can determine which quadrature will be squeezed. If the pump's phase fluctuates, then the quadrature chosen will have a slight admixture[6] of its conjugate quadrature — the noisy quadrature. This argument is treated more carefully in Section II for the case of phase noise in a classical pump. Calculations of the corrections to semiclassical order (*i.e.*, to order $1/\bar{N}$ in the matrix elements, where $\bar{N}$ is the average photon number of the pump) have been previously performed for both the one-[7] and two-mode[8] PA.

Hillery and Zubairy[7] studied the one-mode PA with an interaction Hamiltonian

$$\hat{H}_{\text{int}} = i\frac{\hbar\kappa}{2}(\hat{a}^{\dagger 2}\hat{a}_p - \hat{a}^2\hat{a}_p^{\dagger}) \tag{1.1}$$

(up to a phase rotation of the variables), where $\hat{a}$ and $\hat{a}_p$ are the annihilation operators for the signal and pump modes, respectively, and κ is a coupling constant which is proportional to the second order nonlinear susceptibility $\chi^{(2)}$ of the medium in which the interaction is taking place. They used a path-integral technique[9] to obtain corrections at the semiclassical order. They did not claim to get the full semiclassical correction,[10] and the dominant terms they obtained for the fluctuation in the squeezed quadrature were

$$<\Delta\hat{x}_2^2> \approx \frac{e^{-2u}}{4} + \frac{e^{2u}}{64\bar{N}}, \tag{1.2}$$

where $\hat{x}_2 = -i(\hat{a} - \hat{a}^{\dagger})/2$ is the quadrature (analogous to the position operator), which is squeezed by the interaction in Eq. (1.1), and $u = \bar{N}^{1/2}\kappa t$ is a dimensionless time. This yields a minimum variance, and hence a limit to the squeezing, of

$$<\Delta\hat{x}_2^2>_{\text{min}} \approx \frac{1}{8\bar{N}^{1/2}}, \tag{1.3}$$

which is just what the argument of phase noise in the classical pump gives (see Section II).

Scharf and Walls[8] studied the two-mode PA, whose interaction Hamiltonian is (again up to a rotation of the variables' phases)

$$\hat{H}_{\text{int}} = i\hbar\kappa(\hat{a}_1^{\dagger}\hat{a}_2^{\dagger}\hat{a}_p - \hat{a}_1\hat{a}_2\hat{a}_p^{\dagger}), \tag{1.4}$$

where $\hat{a}_1$ and $\hat{a}_2$ are the annihilation operators for the two signal modes. They used

an asymptotic method developed by Scharf,[11,13] and Scharf and Weiss,[12] to arrive at the dominant correction to the variance of the Hermitian variable

$$\hat{y}_2 \equiv \frac{-i}{2}(\hat{a}_+ - \hat{a}_+^\dagger)\,, \tag{1.5}$$

as

$$<\Delta\hat{y}_2^2> \approx \frac{e^{-2u}}{4} + \frac{e^{6u}}{1920\bar{N}}\,, \tag{1.6}$$

where $\hat{a}_+ \equiv (\hat{a}_1 + \hat{a}_2)/\sqrt{2}$. Scharf and Walls concluded that the minimum variance obtainable by the two-mode PA would be

$$<\Delta\hat{y}_2^2>_{\min} \approx \frac{1}{6(10\bar{N})^{1/4}}\,. \tag{1.7}$$

How can we compare these calculations? If one rewrites Eq. (1.4) in terms of the variables

$$\hat{a}_+ \equiv \frac{1}{\sqrt{2}}(\hat{a}_1 + \hat{a}_2)\,, \tag{1.8a}$$

$$\hat{a}_- \equiv \frac{1}{\sqrt{2}}(\hat{a}_1 - \hat{a}_2)\,, \tag{1.8b}$$

then the interaction Hamiltonian may be written[14]

$$\hat{H}_{\text{int}} = i\frac{\hbar\kappa}{2}(\hat{a}_+^{\dagger 2}\hat{a}_p - \hat{a}_+^2\hat{a}_p^\dagger) - i\frac{\hbar\kappa}{2}(\hat{a}_-^{\dagger 2}\hat{a}_p - \hat{a}_-^2\hat{a}_p^\dagger)\,. \tag{1.9}$$

If the pump is now treated classically, then the $\hat{a}_+$ and $\hat{a}_-$ modes become completely independent, each described by the one-mode PA Hamiltonian Eq. (1.1). Thus, we might expect the same correction to the squeezing that is due to a quantum pump as

found by Hillery and Zubairy [see Eq. (1.2)]. In fact, since the pump is allowed to be quantum mechanical, the $\hat{a}_+$ and $\hat{a}_-$ modes can interact with each other by modifying their common pump. Thus, these modes cannot completely decouple. Even so, Scharf and Walls' results of Eqs. (1.6) and (1.7) are surprising; for a pump with $\bar{N}=10^9$, there is a large discrepancy

$$\frac{<\Delta\hat{y}_2^2>_{\min}}{<\Delta\hat{x}_2^2>_{\min}} = \frac{4}{3}\left[\frac{\bar{N}}{10}\right]^{1/4} = \frac{400}{3} . \tag{1.10}$$

The purpose of this paper is to resolve the apparent discrepancy between these two calculations, first noted by Caves and Crouch.[15] We will use three different methods to calculate the semiclassical corrections for the one- and two-mode PA.

This paper is divided into seven sections. Section II justifies our use of discrete mode calculations for a traveling-wave device, which, in principle, should be given a full continuum treatment, and also reviews the argument for the contribution of phase noise in a classical pump. In Section III, we discuss the symmetries of the discretized PA system and their use in simplifying our calculations.

The first method of calculation (Section IV) involves integrating the Heisenberg equations for the quadrature-phase amplitudes up to the required order. What is novel about the approach presented here is that we first work out the form of all the terms, and then we identify the dominant terms, and proceed to calculate only those terms. This calculation yields the dominant terms up to $O(1/\bar{N}^2)$. By calculating these terms we can estimate the time when the semiclassical correction breaks down. We find that the semiclassical corrections [for instance, Eq. (1.2)] are valid so long as they are much less than one.

The second method (Section V) is numerical. However, we are able to obtain analytic expressions for the full corrections to $O(1/\bar{N}^2)$. We can do this since the general form of these corrections has already been worked out in Section IV. This calculation uses a new algebra, which can be viewed as a semiclassical approximation to the ordinary commutator algebra for annihilation and creation operators. This algebra is used in order to normal-order the annihilation and creation operators of the pump.

The third method (Section VI) uses the positive-P distribution[16] to derive Fokker-Planck equations for one- and two-mode parametric amplifiers. Standard methods of stochastic calculus[17] are then used to derive Ito stochastic differential equations (SDEs) from the Fokker-Planck equations. An approximate solution of the SDEs is obtained by iteration, and the full semiclassical correction is then calculated analytically.

These three methods agree with each other. The latter two show that Hillery and Zubairy have, in fact, calculated the exact semiclassical corrections to the parametric approximation for the one-mode PA, namely,

$$<\Delta\hat{x}_2^2> = \frac{e^{-2u}}{4} + \frac{e^{2u}}{64\bar{N}}[1+(3-8u)e^{-2u} - (1+8u-8u^2)e^{-4u} - 3e^{-6u}]\,. \quad (1.11)$$

Similarly, the exact semiclassical expression for the two-mode PA is found to be

$$<|\Delta\hat{X}_2|^2> = \frac{e^{-2u}}{4} + \frac{e^{2u}}{64\bar{N}}[1+(4-8u)e^{-2u} - (1+12u-8u^2)e^{-4u} - 4e^{-6u}]\,, \quad (1.12)$$

where $X_2 = -i(\hat{a}_1 - \hat{a}_2^\dagger)/2$ is the quadrature-phase operator for the squeezed quadra-

ture, and

$$<|\Delta\hat{X}|^2> = <\hat{X}\hat{X}^\dagger>_{sym} - <\hat{X}><\hat{X}^\dagger>$$

$$= \frac{1}{2}<\hat{X}\hat{X}^\dagger + \hat{X}^\dagger\hat{X}> - <\hat{X}><\hat{X}^\dagger> . \qquad (1.13)$$

The dominant corrections for the one- and two-mode calculations are the same and agree with the dominant correction obtained by Caves and Crouch[15] from a *continuum* calculation.

II. DISCUSSION

The conventional approach to problems in quantum optics typically makes use of a mode expansion to describe the electromagnetic field. Using this approach, one can derive from an appropriate Hamiltonian *temporal* differential equations for the modal creation and annihilation operators, the spatial dependence being carried by the mode functions. Such an approach is suitable for cavity devices in which one has well-defined standing-wave modes (the eigenmodes of the cavity), but not for a traveling-wave device in which such modes are nonexistent. One would like to derive *spatial* differential equations governing the evolution of the field operators through the medium, in analogy with classical nonlinear optics; the conventional approach is clearly unsuited to this purpose. Tucker and Walls[18] and Lane *et al.* [19] recognized these problems with the conventional approach and developed a continuum wave-packet formalism in an attempt to deal with them.

In this section, we briefly describe a discrete mode expansion of the electromagnetic field in terms of *wave-packet* modes that enable us to derive *spatial* equations of motion for PAs. We assume that the wave packets are short compared to the

nonlinear medium through which they propagate, so that they "fit" inside the medium, allowing us to ignore boundary effects. Physically, the individual wave packet propagates from free space through the entrance boundary on a time scale short compared to the time it will spend inside the nonlinear medium; in this way, the interaction is "turned on." This method is preferable to the technique often used in the conventional approach in which the interaction is suddenly turned on throughout all space, either at time $t=0$ or at some time in the remote past. We also present a heuristic argument for the dominant effect of pump quantum fluctuations on the variance of the squeezed quadrature in a PA.

We will give a brief outline of the derivation of the discrete wave-packet mode equations of motion for the PA; details will be given elsewhere. The discrete mode expansions of the signal and pump magnetic field operators in a dispersionless medium are given by

$$\hat{B}_{os}^{(+)}(z,t)=\sum_{n=-M}^{M}\sum_{k=-\infty}^{\infty}\left[\frac{2\pi n_0\hbar\omega_n}{c\sigma T_s}\right]^{1/2}\hat{a}_{nk}(z)\,e^{-i\omega_n[(t-n_0z/c)-kT_s]}$$

$$\times f\,[(t-n_0z/c)-kT_s]\,, \tag{2.1a}$$

$$\hat{B}_{op}^{(+)}(z,t)=i\sum_{k=-\infty}^{\infty}\left[\frac{2\pi n_0\hbar\Omega_p}{c\sigma T_p}\right]^{1/2}\hat{b}_k(z)\,e^{-i\Omega_p[(t-n_0z/c)-kT_p]}$$

$$\times f\,[(t-n_0z/c)-kT_p]\,, \tag{2.1b}$$

where

$$[\hat{a}_{nk}(z),\hat{a}_{n'k'}(z)]=[\hat{b}_k(z),\hat{b}_{k'}(z)]=0\,, \tag{2.2a}$$

$$[\hat{a}_{nk}(z),\hat{a}^{\dagger}_{n'k'}(z)]=\delta_{nn'}\delta_{kk'}, \tag{2.2b}$$

$$[\hat{b}_{k}(z),\hat{b}^{\dagger}_{k'}(z)]=\delta_{kk'}, \tag{2.2c}$$

and

$$f_j(t)=\frac{\sin\pi t/T_j}{\pi t/T_j}, \quad j=s,p \tag{2.3}$$

is the wave-packet envelope function. Here, n_0 is the index of refraction of the dispersionless medium, $\Omega_p=2\Omega$ is the pump frequency, and σ is a cross-sectional area we use to account crudely for the transverse structure of the field. The discrete-mode expansions described by Eqs. (2.1a) and (2.1b) are obtained from a continuum description by dividing the signal and pump bandwidths into "bins" of width $\Delta\omega_s$ and $\Delta\omega_p$, respectively, with signal center frequency $\omega_n=\Omega+n\Delta\omega$ and pump-center frequency $\Omega_p=2\Omega$.[20] Each signal (pump) bin corresponds to a train of wave packets (corresponding to different values of k) in the time domain, each of approximate duration $T_s=2\pi/\Delta\omega_s$ ($T_p=2\pi/\Delta\omega_p$) with envelope given by Eq. (2.3).

By substituting Eqs. (2.1a) and (2.1b) into Maxwell's equations, we obtain the *spatial* equations of motion

$$\frac{d\hat{a}_{nk}(z)}{dz}=\kappa'\sum_{k'=-\infty}^{\infty}e^{i\Omega_p(k'T_p-kT_s)}\frac{\sin[\pi(kT_s-k'T_p)/T_p]}{\pi(kT_s-k'T_p)/T_p}\hat{b}_{k'}(z)\hat{a}^{\dagger}_{-nk}(z), \tag{2.4a}$$

$$\frac{d\hat{b}_{k}(z)}{dz}=-\frac{\kappa'}{2}\sum_{n=-M}^{M}\sum_{k'=-\infty}^{\infty}e^{i\Omega_p(k'T_s-kT_p)}\frac{\sin[\pi(k'T_s-kT_p)/T_p]}{\pi(k'T_s-kT_p)/T_p}$$

$$\times\hat{a}_{nk'}(z)\hat{a}_{-nk'}(z), \tag{2.4b}$$

where the coupling constant κ' is given by

$$\kappa' = \frac{4\pi\Omega\chi^{(2)}}{n_0^2 c}\left[\frac{2\pi n_0 \hbar\Omega_p}{c\sigma T_p}\right]^{1/2} . \tag{2.5}$$

Here, we have assumed that $\Delta\omega_p \ll \Delta\omega_s$ (or $T_p \gg T_s$) to avoid coupling among energy nonconserving modes. By restricting our observations to the region of space-time near $t - n_0 z/c = 0$, we can discard all wave packets (both signal and pump) with $k \neq 0$, since $f_j(kT_j) = \delta_{k,0}$ With $\hat{a}_{n0}(z) \equiv \hat{a}_n(z)$ and $\hat{b}_0(z) \equiv \hat{a}_p(z)$, we find

$$\frac{d\hat{a}_n(z)}{dz} = \kappa' \hat{a}_p(z)\hat{a}^{\dagger}_{-n}(z) , \tag{2.6a}$$

$$\frac{d\hat{a}_p(z)}{dz} = -\frac{\kappa'}{2}\sum_{n=-M}^{M} \hat{a}_n(z)\hat{a}_{-n}(z) . \tag{2.6b}$$

Assuming that the wave packets are narrow compared to the scale of variation set by κ', we can replace z by ct/n_0 and obtain the *temporal* equations of motion

$$\frac{d\hat{a}_n(t)}{dt} = \kappa \hat{a}_p(t)\hat{a}^{\dagger}_{-n}(t) , \tag{2.7a}$$

$$\frac{d\hat{a}_p(t)}{dt} = -\frac{\kappa}{2}\sum_{n=-M}^{M} \hat{a}_n(t)\hat{a}_{-n}(t) , \tag{2.7b}$$

where $\kappa = c\kappa'/n_0$. Equations (2.7a) and (2.7b) are identical to the Heisenberg equations of motion that are derived from the multimode Hamiltonian

$$H = i\frac{\hbar\kappa}{2}\sum_{n=-M}^{M} [\hat{a}_p(t)\hat{a}_n^{\dagger}(t)\hat{a}^{\dagger}_{-n}(t) - \hat{a}_p^{\dagger}(t)\hat{a}_n(t)\hat{a}_{-n}(t)] , \tag{2.8}$$

when the conventional approach is used.

The Hamiltonian Eq. (2.8) correctly describes the interaction of a discrete pump mode with $2M+1$ discrete signal modes, but it does not provide a completely accurate description of traveling-wave parametric amplification, since it ignores the interaction of the pump wave-packet $k=0$ with signal wave packets other than $k=0$. Ignoring as it does interactions with these wave packets, the Hamiltonian cannot correctly describe nonlinear effects such as pump depletion; it does, however, correctly describe the effect of the initial pump quantum fluctuations on the signal modes. We will show, first by a heuristic argument and then by the results of detailed calculations using the Hamiltonian Eq. (2.8), that the initial pump quantum fluctuations are responsible for the dominant correction to the squeezing that is due to the quantum nature of the pump. We also calculate higher-order corrections. By the argument just given, the exact form of these corrections cannot be related to the physical parameters of a traveling-wave PA; these corrections *are* of physical interest, however, in showing how nonlinear effects affect the squeezing, and of mathematical interest in demonstrating the computational tools we have developed to calculate them.

The wave-packet approach gives us a new and more realistic way to deal with traveling-wave problems in quantum optics; it also leads one to realize that the conventional Hamiltonian approach can lead to misleading results when used blindly. We will, however, ignore distinctions between the conventional and the wave-packet approaches through much of this paper. The point we wish to make is that the wave-packet modes are an appropriate set of modes for describing the spatial evolution of quantized electromagnetic fields in traveling-wave devices without resorting to continuum calculations.

We will now give a heuristic argument for the effect of pump quantum fluctuations on the squeezing produced by a parametric amplifier. Our treatment of parametric amplification has thus far treated the pump quantum mechanically. Under certain circumstances, one can treat the pump classically, in what is known as the parametric approximation; in this approximation, one replaces the pump operator by a c-number, $\alpha_p = \bar{N}^{1/2} e^{i\phi_p}$. The interaction Hamiltonian for a one-mode PA, where we ignore all modes in Eq. (2.8) except for $n = 0$, is

$$H_p = i\frac{\hbar\kappa\bar{N}^{1/2}}{2}[\hat{a}^{\dagger 2}(t)\, e^{i\phi_p} - \hat{a}^2(t)\, e^{-i\phi_p}] , \tag{2.9}$$

where $\hat{a}_0(t) \equiv \hat{a}(t)$; two resulting equations of motion are

$$\frac{d\hat{a}(t)}{dt} = \kappa\bar{N}^{1/2} e^{i\phi_p}\, \hat{a}^{\dagger}(t) , \tag{2.10a}$$

$$\frac{d\hat{a}^{\dagger}(t)}{dt} = \kappa\bar{N}^{1/2} e^{-i\phi_p}\, \hat{a}(t) . \tag{2.10b}$$

We define *two* sets of quadrature-phase amplitudes,[14]

$$\hat{x}_1(t) = \frac{1}{2}[\hat{a}(t) + \hat{a}^{\dagger}(t)] , \tag{2.11a}$$

$$\hat{x}_2(t) = -\frac{i}{2}[\hat{a}(t) - \hat{a}^{\dagger}(t)] , \tag{2.11b}$$

and

$$\hat{x}_1'(t) = \frac{1}{2}[\hat{a}(t) e^{-i\phi_p/2} + \hat{a}^{\dagger}(t) e^{i\phi_p/2}] , \tag{2.12a}$$

$$\hat{x}_2'(t) = -\frac{i}{2}[\hat{a}(t) e^{-i\phi_p/2} - \hat{a}^{\dagger}(t) e^{i\phi_p/2}] , \tag{2.12b}$$

the two sets being identical when $\phi_p = 0$. The two sets of quadrature-phase amplitudes are related by the rotation

$$\hat{x}_1(t) = \hat{x}_1'(t)\cos(\phi_p/2) - \hat{x}_2'(t)\sin(\phi_p/2)\,, \tag{2.13a}$$

$$\hat{x}_2(t) = \hat{x}_2'(t)\cos(\phi_p/2) + \hat{x}_1'(t)\sin(\phi_p/2)\,, \tag{2.13b}$$

pictured in Fig. 1 for $\phi_p/2 = \Delta\phi$. By substituting Eqs. (2.12a) and (2.12b) in Eqs. (2.10a) and (2.10b), we see that the quadrature-phase amplitudes $\hat{x}_1'(t)$ and $\hat{x}_2'(t)$ decouple the equations of motion;

$$\frac{d\hat{x}_1'(t)}{dt} = \kappa\bar{N}^{1/2}\hat{x}_1'(t) \quad \Longrightarrow \quad \hat{x}_1'(u) = \hat{x}_1'(0)\,e^{u}\,, \tag{2.14a}$$

$$\frac{d\hat{x}_2'(t)}{dt} = -\kappa\bar{N}^{1/2}\hat{x}_2'(t) \quad \Longrightarrow \quad \hat{x}_2'(u) = \hat{x}_2'(0)\,e^{-u}\,, \tag{2.14b}$$

where $u = \kappa\bar{N}^{1/2}t$ is a dimensionless time. For a vacuum input, one easily finds that

$$<\Delta\hat{x}_1'^2(u)> = \frac{1}{4}e^{2u}\,, \tag{2.15a}$$

$$<\Delta\hat{x}_2'^2(u)> = \frac{1}{4}e^{-2u}\,; \tag{2.15b}$$

the $\hat{x}_2'$ quadrature exhibits maximum squeezing when the pump's phase is ϕ_p. The corresponding noise in the quadrature-phase amplitudes $\hat{x}_1$ and $\hat{x}_2$, from Eqs. (2.13a), (2.13b), (2.15a), and (2.15b) is described by

$$<\Delta\hat{x}_1^2(u)> = \frac{1}{4}e^{2u}\cos^2(\phi_p/2) + \frac{1}{4}e^{-2u}\sin^2(\phi_p/2)\,, \tag{2.16a}$$

$$<\Delta\hat{x}_2^2(u)> = \frac{1}{4}e^{-2u}\cos^2(\phi_p/2) + \frac{1}{4}e^{2u}\sin^2(\phi_p/2)\,. \tag{2.16b}$$

Suppose we allow the pump's phase to fluctuate. For the quantized pump in a coherent state $|\bar{N}^{1/2}>$ with mean photon number $\bar{N}$, the phase fluctuations are characterized by

$$<\Delta\phi_p^2> = <\phi_p^2> = \frac{1}{4\bar{N}}, \tag{2.17}$$

since we may choose without loss of generality $<\phi_p> = 0$. Because $\bar{N}$ is large, $<\Delta\phi_p^2>$ will be small; we can thus approximate $\cos^2(\phi_p/2)$ by 1, $\sin^2(\phi_p/2)$ by $<\phi_p^2/4> = 1/16\bar{N}$, and the variance of the $\hat{x}_2$ quadrature by

$$<\Delta\hat{x}_2^2(u)> \simeq \frac{1}{4}e^{-2u} + \frac{1}{64\bar{N}}e^{2u}. \tag{2.18}$$

The pump can be considered classical and the parametric approximation valid when the correction term is small, that is, when

$$\bar{N} \gg \frac{1}{16}\exp(4\bar{N}^{1/2}\kappa t), \tag{2.19}$$

where we have used the definition of u. The second term of Eq. (2.18) is the dominant correction to the variance of the squeezed quadrature that is due to pump quantum fluctuations. Because of the quantum nature of the pump, phase fluctuations are unavoidable; Figure 1 illustrates their effect. The solid ellipse represents squeezing with a classical pump (i.e., the parametric approximation) with a well-defined phase, $\phi_p = 0$. When $\phi_p \neq 0$, the ellipse is rotated by an angle $\phi_p/2$ as demonstrated by Eqs. (2.13a,b). Pump phase fluctuations cause the orientation of the ellipse to fluctuate about $\phi_p = 0$, with the characteristic angle $\Delta\phi = <\phi_p^2/4> = 1/4\bar{N}^{1/2}$, as represented by the dotted ellipse in Fig. 1, feeding noise from the amplified quadrature into the squeezed quadrature. One also sees why amplitude fluctuations are unimportant.

Amplitude fluctuations merely produce fluctuations in the gain (or rate of squeezing); they do not couple noise in the amplified quadrature into the squeezed quadrature.

The above argument is for a one-mode PA, but it is easily extended to any number of modes; the same argument has been given for a continuum-mode PA,[15] yielding the same dominant correction as given by Eq. (2.18), but with a bandwidth determined by phase mismatching that allows one to relate $\bar{N}$ to the pump power. In the parametric approximation, the signal modes interact in pairs at frequencies $\Omega+n\,\Delta\omega$ and $\Omega-n\,\Delta\omega$; there are no interactions among different pairs in this approximation, and each pair can thus be considered separately. The correction that we have been discussing is due to fluctuations in the initial state of the pump and has nothing to do with back-action from the signal modes — pump depletion being one example of such back action — which would depend on the number of signal modes. The initial fluctuations act on each pair of modes in the same manner as described in Eq. (2.18) for the one-mode PA, yielding *for each pair of modes* a correction identical to that of Eq. (2.18). This correction is then independent of the number of signal modes, justifying our one-mode treatment.

The arguments given above are not rigorous; we have cited quantum mechanics as the ultimate source of pump phase fluctuations, yet we have treated their effect on squeezing classically. What we have given is a plausibility argument for and a physical picture of the dominant effect that is due to such fluctuations. The validity of our arguments will be confirmed by our detailed calculations, showing the correction terms in Eq. (2.18) to be the dominant effect of pump quantum fluctuations on the squeezing, independent of the number of signal modes.

III. SYMMETRIES OF THE MULTIMODE PA

We shall begin our analysis by studying the symmetries of the multimode PA, with the pump in a coherent state and all signal modes initially in vacuum states. This will allow us to concentrate on the few matrix elements that are not constrained by symmetry. Since the multimode PA combines the one- and two-mode PAs as subsystems, we will be able to study the symmetries of these subsystems along with those of the larger system.

From Eqs. (2.7a) and (2.7b), the Heisenberg equations of motion in the interaction picture are

$$\frac{d\hat{a}_n}{dt} = \kappa \hat{a}_p \, \hat{a}^{\dagger}_{-n} \,, \tag{3.1}$$

$$\frac{d\hat{a}_p}{dt} = -\frac{\kappa}{2} \sum_{n=-M}^{M} \hat{a}_n \hat{a}_{-n} \,. \tag{3.2}$$

The quantities that are measured by a balanced homodyne detector are the variances of the quadrature-phase amplitudes,[14]

$$\hat{X}_{(n)1} = \frac{1}{2}(\hat{a}_n + \hat{a}^{\dagger}_{-n}) \,, \tag{3.3a}$$

$$\hat{X}_{(n)2} = -\frac{i}{2}(\hat{a}_n - \hat{a}^{\dagger}_{-n}) \,. \tag{3.3b}$$

Inverting these definitions gives

$$\hat{a}_n = \hat{X}_{(n)1} + i\hat{X}_{(n)2} \,, \tag{3.4a}$$

$$\hat{a}^{\dagger}_{-n} = \hat{X}_{(n)1} - i\hat{X}_{(n)2} \,, \tag{3.4b}$$

and similar definitions for the pump yield

$$\hat{a}_p = \alpha_0 + \hat{P}_1 + i\hat{P}_2 , \tag{3.5}$$

where the coherent amplitude α_0 of the pump has been written explicitly (α_0 is chosen real for convenience). It is worth noting that

$$\hat{X}_{(n)i} = \hat{X}^{\dagger}_{(-n)i} , \tag{3.6}$$

and that $\hat{P}_i$ and $\hat{X}_{(0)i}$ are both Hermitian. The equations of motion [Eqs. (2.7)] in these new variables are

$$\frac{d\hat{X}_{(n)1}}{du} = \hat{X}_{(n)1} + \frac{1}{\alpha_0}(\hat{X}_{(n)1}\hat{P}_1 + \hat{X}_{(n)2}\hat{P}_2) , \tag{3.7a}$$

$$\frac{d\hat{X}_{(n)2}}{du} = -\hat{X}_{(n)2} + \frac{1}{\alpha_0}(\hat{X}_{(n)1}\hat{P}_2 - \hat{X}_{(n)2}\hat{P}_1) , \tag{3.7b}$$

$$\frac{d\hat{P}_1}{du} = -\frac{1}{2\alpha_0}\sum_{n=-M}^{M}(\hat{X}_{(n)1}\hat{X}^{\dagger}_{(n)1} - \hat{X}_{(n)2}\hat{X}^{\dagger}_{(n)2}) , \tag{3.7c}$$

$$\frac{d\hat{P}_2}{du} = -\frac{1}{2\alpha_0}\sum_{n=-M}^{M}(\hat{X}_{(n)1}\hat{X}^{\dagger}_{(n)2} + \hat{X}_{(n)2}\hat{X}^{\dagger}_{(n)1}) , \tag{3.7d}$$

where $u \equiv \kappa\alpha_0 t$.

We are interested in symmetries of the time evolved matrix elements. Thus, we want symmetries that preserve both the equations of motion and the initial state of the system. The coherent part of the pump has been subtracted out in our choice of variables [see Eq. (3.5)]; in terms of the pump quadrature-phase amplitudes $\hat{P}_1$ and $\hat{P}_2$, the effective initial state of the pump is vacuum. As any phase shift, reflection, or rotation of the quadratures leaves the vacuum invariant, we seek such transformations

which leave the equations of motion [Eq. (3.7)] invariant.

The symmetries may now be classed as follows.

i) Time-reversal:

$$\hat{X}_{(n)1} \leftrightarrow \hat{X}_{(n)2} , \tag{3.8a}$$

$$\hat{P}_2 \rightarrow -\hat{P}_2 , \tag{3.8b}$$

$$u \rightarrow -u . \tag{3.8c}$$

One consequence of this symmetry is that the squeezed quadrature is simply the time-reversed amplified quadrature:

$$<|\Delta\hat{X}_{(n)1}(u)|^2> = <|\Delta\hat{X}_{(n)2}(-u)|^2> . \tag{3.9}$$

ii) Reflection:

$$\hat{X}_{(n)1} \rightarrow -\hat{X}_{(n)1} \qquad (\text{or } \hat{X}_{(n)2} \rightarrow -\hat{X}_{(n)2}) , \tag{3.10a}$$

$$\hat{P}_2 \rightarrow -\hat{P}_2 , \tag{3.10b}$$

which tells us that

$$<\hat{X}_{(n)1}(u)> = <\hat{X}_{(n)2}(u)> = <\hat{P}_2(u)> \equiv 0 . \tag{3.11}$$

Hence, the pump's phase does not drift, and the signal modes do not acquire a coherent piece. A similar symmetry for the pump amplitude $\hat{P}_1$ is absent because the pump may, for example, become depleted. Another consequence of this symmetry is that the conjugate quadratures remain uncorrelated; *e.g.*,

$$<\hat{X}_{(n)1}(u)\hat{X}^{\dagger}_{(n)2}(u)> \equiv 0 . \tag{3.12}$$

iii) Time stationary noise:[14]

$$\hat{X}_{(n)1}(u)\rightarrow e^{i\theta_n}\hat{X}_{(n)1}(u)\,, \qquad n\neq 0\,, \tag{3.13a}$$

$$\hat{X}_{(n)2}(u)\rightarrow e^{i\theta_n}\hat{X}_{(n)2}(u)\,, \qquad n\neq 0\,. \tag{3.13b}$$

We cannot apply this symmetry to the one-mode PA (corresponding to $n=0$), since the quadratures $\hat{X}_{(0)1}$ and $\hat{X}_{(0)2}$ are Hermitian [see Eqs. (3.3)]. But we can say that matrix elements, like $<\hat{X}_{(n)i}(u)\hat{X}_{(n')j}(u)>$ and $<\hat{X}_{(n)i}(u)\hat{X}_{(n')j}(u)\hat{X}^{\dagger}_{(m)k}(u)>$ are identically zero for $n\neq 0$. With a combination of the above symmetries we may determine all quadratic matrix elements given only

$$<|\Delta\hat{X}_{(n)2}|^2>=<\hat{X}_{(n)2}\hat{X}^{\dagger}_{(n)2}>_{\rm sym}\,. \tag{3.14}$$

At this point it is worth asking how the quantity $<\Delta\hat{y}_2^2>$ [Eq. (1.6)] calculated by Scharf and Walls[8] is related to the quadrature-phase amplitudes. They considered the two-mode PA and the quantity

$$\hat{y}_2=\frac{1}{\sqrt{2}}(\hat{X}_{(1)2}+\hat{X}^{\dagger}_{(1)2}) \tag{3.15}$$

[Eqs. (1.5), (1.8), and (3.3)]. When the time stationary noise and reflection symmetries are used, one finds

$$\begin{aligned}<\Delta\hat{y}_2^2>&=\frac{1}{2}<\hat{X}^2_{(1)2}>+\frac{1}{2}<\hat{X}^{\dagger}_{(1)2}>+<\hat{X}_{(1)2}\hat{X}^{\dagger}_{(1)2}>_{\rm sym}-<\hat{y}_2>^2\\ &=<|\Delta\hat{X}_{(1)2}|^2>\,,\end{aligned} \tag{3.16}$$

which is just the variance of the squeezed quadrature-phase amplitude.

IV. DOMINANT TERMS FOR THE MULTIMODE PA

In this section, we shall determine the dominant time behavior of the squeezed quadrature-phase amplitude that is due to the quantum nature of the pump. As in Section III, we study the multimode PA, extracting the one- and two-mode results at the end.

The clear way to proceed in obtaining an expansion in α_0^{-1} is to take the Heisenberg equations [Eq. (3.7)] and iterate them. It is easier first to write them as integral equations:

$$\hat{X}_{(n)1}(u)=e^{u}\hat{X}_{(n)1}(0)+\frac{e^{u}}{\alpha_0}\int_0^u du' e^{-u'}\left[\hat{X}_{(n)1}(u')\hat{P}_1(u')+\hat{X}_{(n)2}(u')\hat{P}_2(u')\right] , \quad (4.1a)$$

$$\hat{X}_{(n)2}(u)=e^{-u}\hat{X}_{(n)2}(0)+\frac{e^{-u}}{\alpha_0}\int_0^u du' e^{u'}\left[\hat{X}_{(n)1}(u')\hat{P}_2(u')-\hat{X}_{(n)2}(u')\hat{P}_1(u')\right] \quad (4.1b)$$

$$\hat{P}_1(u)=\hat{P}_1(0)-\frac{1}{2\alpha_0}\int_0^u du' \sum_{n=-M}^{M}\left[\hat{X}_{(n)1}(u')\hat{X}^{\dagger}_{(n)1}(u')-\hat{X}_{(n)2}(u')\hat{X}^{\dagger}_{(n)2}(u')\right] , \quad (4.1c)$$

$$\hat{P}_2(u)=\hat{P}_2(0)-\frac{1}{2\alpha_0}\int_0^u du' \sum_{n=-M}^{M}\left[\hat{X}_{(n)1}(u')\hat{X}^{\dagger}_{(n)2}(u')+\hat{X}_{(n)2}(u')\hat{X}^{\dagger}_{(n)1}(u')\right] . \quad (4.1d)$$

By substituting the quadrature-phase amplitudes correct to $O(1/\alpha_0^n)$ into Eq. (4.1), we will obtain expressions for them correct to $O(1/\alpha_0^{n+1})$. Although this procedure is easy up to $O(1/\bar{N})$ [*i.e.*, $O(1/\alpha_0^2)$], it becomes prohibitive in obtaining even the $O(1/\bar{N}^2)$ correction, this correction requiring around 1000 terms. Instead of keeping every term we shall determine which terms can yield a dominant contribution and then calculate them.

Our objective is to determine the time dependence of the various terms that appear at each order in the expansion. Thus, we start by keeping only information

about time dependence in the integral equations (4.1). This leaves us with the ''equations''

$$X_1(u) \simeq e^u + \frac{e^u}{\alpha_0}\int_0^u du' e^{-u'}\Big[X_1(u')P_1(u') + X_2(u')P_2(u')\Big] \ , \tag{4.2a}$$

$$X_2(u) \simeq e^{-u} + \frac{e^{-u}}{\alpha_0}\int_0^u du' e^{u'}\Big[X_1(u')P_2(u') + X_2(u')P_1(u')\Big] \ , \tag{4.2b}$$

$$P_1(u) \simeq 1 + \frac{1}{\alpha_0}\int_0^u du'\Big[X_1(u')X_1(u') + X_2(u')X_2(u')\Big] \ , \tag{4.2c}$$

$$P_2(u) \simeq 1 + \frac{1}{\alpha_0}\int_0^u du'\Big[X_1(u')X_2(u') + X_2(u')X_1(u')\Big] \ , \tag{4.2d}$$

where we have thrown away numerical coefficients and sums over different modes. We shall even treat these equations as c-number equations. Next, we define the symbol

$$\tilde{u}^n \equiv u^n + u^{n-1} + \cdots + u + 1 \tag{4.3}$$

to represent an arbitrary polynomial of order n in the scaled time with all its coefficients suppressed. With these simplifications, Eqs. (4.2) become easy to iterate. The information we are left with is the *form* of solution at each order in the expansion; for instance, the squeezed quadrature-phase amplitude has the form

$$\begin{aligned} X_2 \simeq e^{-u} &+ \frac{1}{\alpha_0}(e^u + \tilde{u}e^{-u}) + \frac{1}{\alpha_0^2}(e^u + \tilde{u}^2 e^{-u} + e^{-3u}) \\ &+ \frac{1}{\alpha_0^3}(e^{3u} + \tilde{u}^2 e^u + \tilde{u}^3 e^{-u} + \tilde{u}e^{-3u}) \\ &+ \frac{1}{\alpha_0^4}(\tilde{u}e^{3u} + \tilde{u}^3 e^u + \tilde{u}^4 e^{-u} + \tilde{u}^2 e^{-3u} + e^{-5u}) + \cdots \end{aligned} \tag{4.4}$$

to $O(1/\alpha_0^4)$. Squaring this expression gives us the form of the variance:

$$<X_2^2> \approx e^{-2u} + \frac{1}{\bar{N}}(e^{2u} + \tilde{u} + \tilde{u}^2 e^{-2u} + e^{-4u})$$

$$+ \frac{1}{\bar{N}^2}(e^{4u} + \tilde{u}^2 e^{2u} + \tilde{u}^3 + \tilde{u}^4 e^{-2u} + \tilde{u}^2 e^{-4u} + e^{-6u}) + \cdots . \quad (4.5)$$

The terms multiplied by an odd power of α_0^{-1} contain an odd number of quadrature operators and thus vanish by the reflection symmetry. The dominant terms at each order in $1/\bar{N}$ are pure exponentials. Thus, when we calculate the dominant terms for the squeezed quadrature, we may throw away any polynomial times an exponential, if the polynomial is nontrivial.

We now use these simplifications to iterate Eq. (4.1), retaining only the dominant term at each order:

$$\hat{X}_{(n)1}(u) \approx e^u \hat{X}_{(n)1}(0) - \frac{e^{3u}}{8\alpha_0^2}\hat{X}_{(n)1}(0) \sum_{m=-M}^{M} \hat{X}_{(m)1}(0)\hat{X}^{\dagger}_{(m)1}(0) + O(1/\alpha_0^3) , \quad (4.6a)$$

$$\hat{X}_{(n)2}(u) \approx e^{-u}\hat{X}_{(n)2}(0) + \frac{e^u}{2\alpha_0}\hat{X}_{(n)1}(0)\hat{P}_2(0)$$

$$- \frac{e^{3u}}{16\alpha_0^3}\hat{X}_{(n)1}(0) \sum_{m=-M}^{M} \hat{X}_{(m)1}(0)\hat{X}^{\dagger}_{(m)1}(0)\hat{P}_2(0)$$

$$+ O(1/\alpha_0^4) , \quad (4.6b)$$

$$\hat{P}_1(u) \approx \hat{P}_1(0) - \frac{e^{2u}}{4\alpha_0} \sum_{m=-M}^{M} \hat{X}_{(m)1}(0)\hat{X}^{\dagger}_{(m)1}(0) + O(1/\alpha_0^3) , \quad (4.6c)$$

$$\hat{P}_2(u) \approx \hat{P}_2(0) - \frac{e^{2u}}{4\alpha_0^2} \sum_{m=-M}^{M} \hat{X}_{(m)1}(0)\hat{X}^{\dagger}_{(m)1}(0)\hat{P}_2(0) + O(1/\alpha_0^3) . \quad (4.6d)$$

This gives the variance for the quadrature $\hat{X}_{(n)2}$ as

$$<|\Delta\hat{X}_{(n)2}(u)|^2> \approx e^{-2u}<\hat{X}_{(n)2}(0)\hat{X}^{\dagger}_{(n)2}(0)>_{\text{sym}}$$

$$+\frac{e^{2u}}{4\bar{N}}<\hat{X}_{(n)1}(0)\hat{X}^{\dagger}_{(n)1}(0)>_{\text{sym}}<\hat{P}^2_2(0)>$$

$$-\frac{e^{4u}}{64\bar{N}^2}<2\hat{X}_{(n)1}(0)\sum_m\hat{X}_{(m)1}(0)\hat{X}^{\dagger}_{(m)1}(0)\hat{X}^{\dagger}_{(n)1}(0)$$

$$+\hat{X}^{\dagger}_{(n)1}(0)\hat{X}_{(n)1}(0)\sum_m\hat{X}_{(n)1}(0)\hat{X}^{\dagger}_{(n)1}(0)$$

$$+\sum_m\hat{X}_{(n)1}(0)\hat{X}^{\dagger}_{(n)1}(0)\hat{X}^{\dagger}_{(n)1}(0)\hat{X}_{(n)1}(0)>$$

$$<\hat{P}^2_2(0)>+O(1/\bar{N}^3)\,. \tag{4.7}$$

Restricting this calculation to one mode ($M=0$; $n=0,$, only) gives

$$<\Delta\hat{x}^2_2(u)> \approx \frac{e^{-2u}}{4}+\frac{e^{2u}}{64\bar{N}}-\frac{3e^{4u}}{1024\bar{N}^2}\,, \tag{4.8}$$

and for two modes ($n=\pm1$, only)

$$<|\Delta\hat{X}_2(u)|^2> \approx \frac{e^{-2u}}{4}+\frac{e^{2u}}{64\bar{N}}-\frac{4e^{4u}}{1024\bar{N}^2}\,. \tag{4.9}$$

When $M \geq 0$, the full M-mode calculation gives

$$<|\Delta\hat{X}_{(n)2}(u)|^2> \approx \frac{e^{-2u}}{4}+\frac{e^{2u}}{64\bar{N}}-\frac{(3+4M)e^{4u}}{1024\bar{N}^2}\,, \tag{4.10}$$

which confirms our argument that the dominant semiclassical correction is

independent of the number of signal modes.

For the M-mode PA the dominant term at $O(1/\bar{N}^2)$ (and higher orders) is small compared to the $O(1/\bar{N})$ term when

$$e^{2u} \ll \frac{16\bar{N}}{3+4M} . \tag{4.11}$$

This is similar to the restrictions for the one-mode PA (*i.e.*, $M=0$) calculations done by Hillery and Zubairy,[7] except that they need the added restriction on their results that $u \lesssim 1$.

Our result shows that so long as the condition in Eq. (4.11) holds, the semiclassical approximation is sufficient to determine the limit to the squeezing, and thus the squeezing is limited by the pump phase-noise. If the condition of Eq. (4.11) does not hold, the semiclassical approximation breaks down, and only a full quantum treatment to all orders can be relied upon.

V. SEMI-NUMERICAL METHOD

The central theme of this paper is to determine the behavior of the squeezed quadrature's variance $<|\Delta\hat{X}_2|^2>$ as a function of the scaled time u. Looking at Eq. (4.5), we can see that we have already derived the form of this variance to $O(1/\bar{N}^2)$, and in Eqs. (4.8) to (4.10) we have also calculated the coefficients of the dominant terms to this order. It is nonetheless worth calculating the coefficients for the sub-dominant terms appearing in Eq. (4.5), since this allows us to check directly whether or not these terms have coefficients large enough to overcome their smaller relative growth at times when the phase noise begins to dominate. Further, we shall show that it is relatively simple to derive the exact expression for $<|\Delta\hat{X}_2|^2>$ to $O(1/\bar{N}^2)$ (or

higher), using a new algebra and some numerical assistance. We hope that exposition of this new technique will be an adequate motivation for presenting the exact form of $<|\Delta\hat{X}_2|^2>$ to $O(1/\bar{N}^2)$ for the one- and two-mode PA. In Section II we showed how to discretize a continuum mode calculation. Clearly, the fineness of this discretization should not appear in any physical quantities relevant to the continuum system. We find, as we argued in Section II, that all but the dominant correction at $O(1/\bar{N})$ do depend in detail on the discretization prescription. If ever these terms become important, then the discrete mode equations no longer model accurately the continuum system.

Let us see how we might proceed in finding the unknown coefficients in Eq. (4.5). To $O(\bar{N}^0)$, there is one coefficient; to $O(1/\bar{N}^1)$, there are seven; and to $O(1/\bar{N}^2)$ there are seventeen. We can calculate the first terms in a power series expansion of $<|\Delta\hat{X}_2|^2>$ in the scaled time u; similarly, we can do a power-series expansion of the form given in Eq. (4.5) with a set of unknown coeffients. By equating the coefficients of u and $\bar{N}$ in these expansions, we can obtain sufficient simultaneous equations to solve for the unknown coefficients of Eq. (4.5).

Since at $O(1/\bar{N}^2)$ we must find seventeen unknown coefficients, we will need an expansion of $<|\Delta\hat{X}_2|^2>=<|\hat{X}_2|^2>$ to $O(u^{16})$. To avoid repetition, we shall describe the calculation only for the one-mode PA with interaction Hamiltonian $\hat{H}_{\rm int}=i\kappa(a^{\dagger 2}b-a^2b^{\dagger})/2$; the variance of the evolving quadrature-phase amplitude

may be written

$$<0;\alpha_0|\hat{x}_2^2(u)|0;\alpha_0> = <0;\alpha_0|\exp(iu\hat{H}_{\rm int}/\kappa\alpha_0)\hat{x}_2^2\exp(-iu\hat{H}_{\rm int}/\kappa\alpha_0)|0;\alpha_0>$$

$$= <0;\alpha_0|\hat{x}_2^2 + iu\left[\frac{\hat{H}_{\rm int}}{\kappa\alpha_0},\hat{x}_2^2\right] - \frac{u^2}{2}\left[\frac{\hat{H}_{\rm int}}{\kappa\alpha_0},\left[\frac{\hat{H}_{\rm int}}{\kappa\alpha_0},\hat{x}_2^2\right]\right]$$

$$+ \cdots |0;\alpha_0> , \qquad (5.1)$$

where the scaled time is $u = \kappa\alpha_0 t$, and $|0;\alpha_0> = |0>\otimes|\alpha_0>$ is the initial state (vacuum for the signal mode and coherent state for the pump mode). The signal and pump mode annihilation operators are a and b, respectively. Since the signal mode is in vacuum its annihilation and creation operators may be treated directly as ladder operators on a number state. For the pump mode, we need to normal-order its annihilation and creation operators.

Our first simplification comes from needing only the next to semiclassical approximation; *i.e.*, $O(1/\bar{N}^2)$, for the calculation. Thus, when we are normal-ordering the pump mode operators, we may throw away all terms generated by more than two applications of the commutation relation $[b, b^\dagger] = 1$. This can be done automatically by using a new algebra, which we now present. For simplicity, we start with a description of this algebra good up to semiclassical order [*i.e.*, $O(1/\bar{N})$].

Let us start with more general considerations: Many calculations in quantum optics require the expectation value of a product of creation and annihilation operators in a coherent state; *e.g.*,

$$<\beta|b^{\dagger 3}b^2b^\dagger b^3b^\dagger b|\beta> . \qquad (5.2)$$

If we are interested only in calculating this to semiclassical order, then we need only

keep the terms up to $O(1/|\beta|^2)$ times the dominant (''classical'') term. The dominant term is given by replacing the operators b and $b^\dagger$ with the c-numbers β and β^*, respectively. For the semiclassical correction, we need only perform the first of a series of commutation relations – never performing more than one for each term [since we are interested only in terms to order $\hbar$ or equivalently $O(1/|\beta|^2)$ times the classical piece]. To do this, we make use of the commutation relation

$$[b, b^\dagger] = 1 \tag{5.3}$$

and the notation

$$B^{\sigma}_{n,m} \equiv b^{\dagger n} b^m + \sigma b^{\dagger n-1} b^{m-1}\,; \tag{5.4}$$

the creation and annihilation operators are written as

$$B^{0}_{1,0} = b^\dagger\,, \tag{5.5a}$$

$$B^{0}_{0,1} = b\,. \tag{5.5b}$$

To semiclassical order in the amplitude $|\beta|$, the following relation allows us to normal-order any expansion:

$$B^{\sigma}_{n,m} B^{\sigma'}_{n',m'} = B^{\sigma+\sigma'+mn'}_{n+n',m+m'}\,. \tag{5.6}$$

Clearly, $m \times n'$ is just the number of times the commutation relation is required in order to pass m annihilation operators from $B^{\sigma}_{n,m}$ past the n' creation operators in $B^{\sigma'}_{n',m'}$. As an example, the evaluation of the matrix element in Eq. (5.2) can be

worked out to semiclassical order in the coherent state of complex amplitude β:

$$<\beta|b^{\dagger 3}b^2b^\dagger b^3b^\dagger b|\beta> = <\beta|B^{\,0}_{3,2}B^{\,0}_{1,3}B^{\,0}_{1,1}|\beta>$$

$$= <\beta|B^{\,7}_{5,6}|\beta> = \beta^{*5}\beta^6 + 7\beta^{*4}\beta^5 + O(|\beta|^7)\,. \qquad (5.7)$$

For calculations at next to semiclassical order, we simply extend the algebra to be good up to $O(1/\bar{N}^2)$. In this case, we modify the notation so

$$B^{\,\tau}_{n,m} \equiv b^{\dagger n}b^m + \sigma b^{\dagger n-1}b^{m-1} + \tau b^{\dagger n-2}b^{m-2}\,, \qquad (5.8)$$

$$B^{\,0}_{1,0} = b^\dagger\,, \qquad (5.9a)$$

$$B^{\,0}_{0,1} = b\,. \qquad (5.9b)$$

After some calculation we find

$$B^{\,\tau}_{n,m}B^{\,\tau'}_{n',m'} = B^{\,\tau+\tau'+\sigma\sigma'+\sigma(m-1)n'+\sigma' m(n'-1)+(m,n')}_{n+n',m+m'}\,, \qquad (5.10)$$

where

$$(m,n) = m(m-1)n(n-1)/2\,. \qquad (5.11)$$

This extended algebra allows us, for example, to determine the $O(|\beta|^2)$ term in Eq. (5.7) to be $8\beta^{*3}\beta^4$.

In calculating $<\hat{x}_2^2>$, rather than proceeding precisely as suggested by Eq. (5.1), we used this "B-algebra" to obtain, instead, the time evolved state of the sys-

tem

$$\exp[u(a^{\dagger 2}b - a^2 b^\dagger)/2]\,|0;\alpha_0\rangle = |0;\alpha_0\rangle + \frac{\sqrt{2}u}{2} B^0_{0,1}\,|2;\alpha_0\rangle$$

$$+ \frac{u^2}{8}(\sqrt{4!}B^0_{0,2}\,|4;\alpha_0\rangle - 2B^0_{1,1}\,|0;\alpha_0\rangle)$$

$$+ \cdots . \qquad (5.12)$$

Here, $|n;\alpha_0\rangle = |n\rangle \otimes |\alpha_0\rangle$ is the outer product of a number state for the signal with the pump's initial coherent state. Equation (5.12) shows us the first terms in a power-series expansion in u, although we actually retained the first 17 terms. To get $\langle \mathcal{X}_2^2 \rangle$, we need only wedge the operator $\mathcal{X}_2^2$ between pairs of the time-evolved states given by Eq. (5.12).

Even after these simplifications, the calculation would still be very tedious to do by hand, so we calculated Eq. (5.12) on a computer. The evolved state was represented by a multidimensional array, and numbers were calculated with limited accuracy. This allowed us to generate the power series expansion of $\langle \mathcal{X}_2^2 \rangle$ and hence, the coefficients of Eq. (4.5) up to the accuracy used in the computations.

The final step comes in estimating the accuracy necessary to reproduce the rational coefficients that should appear in Eq. (4.5). That they are indeed rational can be seen by looking at how they would arise if we were to iterate the Heisenberg equations [Eq. (4.1)] in full. This also allows us to estimate an upper bound on the numerators and denominators for each fraction. In the worst case, to $O(1/\bar{N})$, the numerator could come from all of the sixteen terms that appear at that order, and the denominator from the factors of two in the definition of the quadrature phase

amplitudes and from factors that are due to the time integrations. An unambiguous calculation of this worst case rational number requires only about four significant figures in the final answer. Similarly, at $O(1/\bar{N}^2)$, we need to keep ten significant figures in the final coefficients.

For the one-mode PA, we find at next to semiclassical order that

$$<\Delta\hat{x}_2^2> = \text{[Eq.(1.11)]} - \frac{1}{1024\bar{N}^2}\Big[3e^{4u} - (19/4 - 24u + 32u^2)e^{2u} + 60 - 112u$$

$$- (80 + 58u - 48u^2 - 128u^3/3 + 32u^4)e^{-2u}$$

$$+ (33 + 48u + 96u^2)e^{-4u} - 45e^{-6u}/4\Big] \quad .(5.13)$$

For the two-mode PA the same technique yields a next to semiclassical result

$$<|\Delta\hat{X}_2|^2> = \text{[Eq.(1.12)]} - \frac{1}{1024\bar{N}^2}\Big[4e^{4u} - (5 - 28u + 40u^2)e^{2u} + 96 - 160u$$

$$- (105 + 112u - 48u^2 - 224u^3/3 + 32u^4)e^{-2u}$$

$$+ (28 + 32u + 128u^2)e^{-4u} - 18e^{-6u}\Big] \quad . \qquad (5.14)$$

VI. STOCHASTIC DIFFERENTIAL EQUATIONS

A. The one-mode PA

The dynamic evolution of a one-mode PA is described by von Neumann's equation in the interaction picture

$$\frac{\partial \hat{\rho}_I(t)}{\partial t} = \frac{i}{\hbar}[\hat{\rho}_I(t), \hat{H}_{\text{int}}(t)] , \tag{6.1}$$

where the interaction Hamiltonian $H_{\text{int}}(t)$ is given by Eq. (1.1). All operators are now in the interaction picture. We will assume that initially the signal mode is in the vacuum state, and the pump is in a coherent state of real amplitude α_0:

$$\hat{\rho}_I(0) = |0;\alpha_0\rangle\langle 0;\alpha_0| . \tag{6.2}$$

To solve Eq. (6.1), it is convenient to project the density operator $\hat{\rho}_I(t)$ onto a suitable set of basis states. The positive-P representation[16] is an off-diagonal representation obtained from an expansion on a coherent state basis:

$$\hat{\rho}_I(t) = \iiiint P(\alpha,\alpha_p,\beta,\beta_p,t)\,\hat{\Lambda}(\alpha,\alpha_p,\beta,\beta_p)\,d^2\alpha\, d^2\alpha_p\, d^2\beta\, d^2\beta_p , \tag{6.3}$$

where the operator $\hat{\Lambda}$ is given by

$$\hat{\Lambda}(\alpha,\alpha_p,\beta,\beta_p) \equiv \frac{|\alpha;\alpha_p\rangle\langle\beta^*;\beta_p^*|}{\langle\beta^*;\beta_p^*|\alpha;\alpha_p\rangle}$$

$$= e^{-(\alpha\beta+\alpha_p\beta_p)} e^{\alpha\hat{a}^\dagger+\alpha_p\hat{a}_p^\dagger}|0;0\rangle\langle 0;0|e^{\beta\hat{a}+\beta_p\hat{a}_p} . \tag{6.4}$$

By substituting Eqs. (1.1), (6.3), and (6.4) into Eq. (6.1) and integrating by parts, we

find the Fokker-Planck equation

$$\frac{\partial P}{\partial \tau} = \left[-\alpha_p \beta \frac{\partial}{\partial \alpha} - \alpha \beta_p \frac{\partial}{\partial \beta} + \frac{\alpha^2}{2} \frac{\partial}{\partial \alpha_p} + \frac{\beta^2}{2} \frac{\partial}{\partial \beta_p} + \frac{\alpha_p}{2} \frac{\partial^2}{\partial \alpha^2} + \frac{\beta_p}{2} \frac{\partial^2}{\partial \beta^2} \right] P \ , \tag{6.5}$$

where $\tau = \kappa t$ and $P \equiv P(\alpha, \alpha_p, \beta, \beta_p, \tau)$.

In its present form, Eq. (6.5) is a complex, eight-dimensional Fokker-Planck equation. The analyticity of $\hat{\Lambda} \equiv \hat{\Lambda}(\alpha, \alpha_p, \beta, \beta_p)$, however, allows us some freedom of choice in interpreting the derivatives.[21] By properly interpreting the derivatives in Eq. (6.5), we obtain a real Fokker-Planck equation with positive-semidefinite diffusion. Using the standard methods of stochastic calculus,[17] this eight-dimensional Fokker-Planck equation yields a set of eight real, first-order Ito stochastic differential equations (SDE's). When written in complex notation, the resulting SDE's are

$$d\alpha = \alpha_p \beta \, d\tau + \sqrt{\alpha_p} \, dW_1 \ , \tag{6.6a}$$

$$d\beta = \alpha \beta_p \, d\tau + \sqrt{\beta_p} \, dW_2 \ , \tag{6.6b}$$

$$d\alpha_p = -\frac{1}{2} \alpha^2 \, d\tau \ , \tag{6.6c}$$

$$d\beta_p = -\frac{1}{2} \beta^2 \, d\tau \ , \tag{6.6d}$$

where, in the Ito calculus,

$$dW_1^2 = dW_2^2 = d\tau \ . \tag{6.6e}$$

The Wiener increments dW_1 and dW_2 are real and independent.

Although we cannot solve Eqs. (6.6) analytically, an approximate solution is possible for a nearly classical pump. We assume that the stochastic pump mode variables α_p and β_p consist of a mean amplitude α_0 (chosen to be real) plus fluctuations $\Delta\alpha$ and $\Delta\beta$:

$$\alpha_p \equiv \alpha_0 + \Delta\alpha \,, \tag{6.7a}$$

$$\beta_p \equiv \alpha_0 + \Delta\beta \,. \tag{6.7b}$$

We define new variables x_1, p_1, x_2, and p_2 by

$$x_1 = \frac{1}{2}(\alpha+\beta) \,, \qquad x_2 = -\frac{i}{2}(\alpha-\beta) \,, \tag{6.8a}$$

$$p_1 = \frac{1}{2}(\Delta\alpha+\Delta\beta) \,, \quad p_2 = -\frac{i}{2}(\Delta\alpha-\Delta\beta) \,. \tag{6.8b}$$

It is convenient to change variables once again. We define the variables z_1 and z_2 by

$$z_1 = x_1 e^{-u} \,, \tag{6.9a}$$

$$z_2 = x_2 e^{u} \,, \tag{6.9b}$$

where $u = \alpha_0 \tau$. The resulting SDE's are

$$dz_1 = \frac{1}{\alpha_0}(z_1 p_1 + z_2 p_2 e^{-2u})\, du + \frac{1}{2} e^{-u} \left[\left(1 + \frac{p_1 + i p_2}{\alpha_0} \right)^{1/2} dV_1 \right.$$

$$\left. + \left(1 + \frac{p_1 - i p_2}{\alpha_0} \right)^{1/2} dV_2 \right] \,, \tag{6.10a}$$

$$dz_2 = \frac{1}{\alpha_0}(z_1 p_2 e^{2u} - z_2 p_1)\,du - \frac{i}{2}e^{u}\left[\left(1+\frac{p_1+ip_2}{\alpha_0}\right)^{1/2} dV_1 - \left(1+\frac{p_1-ip_2}{\alpha_0}\right)^{1/2} dV_2\right], \tag{6.10b}$$

$$dp_1 = \frac{1}{2\alpha_0}(z_2^2 e^{-2u} - z_1^2 e^{2u})\,du\ , \tag{6.10c}$$

$$dp_2 = -\frac{1}{\alpha_0} z_1 z_2\,du\ , \tag{6.10d}$$

where $dV_1 = \sqrt{\alpha_0}\,dW_1$ and $dV_2 = \sqrt{\alpha_0}\,dW_2$.

We use an iterative procedure to obtain an approximate solution of Eqs. (6.10). The square roots are expanded in a Taylor series:

$$\left(1+\frac{p_1 \pm ip_2}{\alpha_0}\right)^{1/2} = 1 + \frac{1}{\alpha_0}\frac{p_1 \pm ip_2}{2} - \frac{1}{\alpha_0^2}\frac{(p_1 \pm ip_2)^2}{8} + \cdots\ . \tag{6.11}$$

Substituting Eq. (6.11) into Eqs. (6.10) and integrating formally, we find

$$z_1(u) = z_1(0) + \frac{1}{\alpha_0}\int_0^u [z_1(x)p_1(x) + z_2(x)p_2(x)e^{-2x}]\,dx + \frac{1}{\sqrt{2}}\int_0^u e^{-x}\left\{dV + \frac{1}{\alpha_0}\left[\frac{1}{2}p_1(x)dV + \frac{i}{2}p_2(x)dW\right] + \cdots\right\}, \tag{6.12a}$$

$$z_2(u) = z_2(0) + \frac{1}{\alpha_0}\int_0^u [z_1(x)p_2(x)e^{2x} - z_2(x)p_1(x)]\,dx - \frac{i}{\sqrt{2}}\int_0^u e^{x}\left\{dW + \frac{1}{\alpha_0}\left[\frac{1}{2}p_1(x)dW + \frac{i}{2}p_2(x)dV\right] + \cdots\right\}, \tag{6.12b}$$

$$p_1(u)=p_1(0)+\frac{1}{2\alpha_0}\int_0^u [z_2^2(x)e^{-2x}-z_1^2(x)e^{2x}]\,dx\ , \tag{6.12c}$$

$$p_2(u)=p_2(0)-\frac{1}{\alpha_0}\int_0^u z_1(x)z_2(x)\,dx\ . \tag{6.12d}$$

Here, we have defined two new independent Wiener increments

$$dV(x)=\frac{dV_1(x)+dV_2(x)}{\sqrt{2}}\ , \tag{6.13a}$$

$$dW(x)=\frac{dV_1(x)-dV_2(x)}{\sqrt{2}}\ . \tag{6.13b}$$

The new Wiener increments defined in Eqs. (6.13) correspond to a rotation of the old Wiener increments, dV_1 and dV_2, and hence retain the same correlation matrix.[17]

We can ignore the initial values $z_1(0)=x_1(0)$, $z_2(0)=x_2(0)$, $p_1(0)$ and $p_2(0)$ in subsequent calculations, because all moments involving these quantities are zero. To see this, we observe that the P-function gives normally ordered averages for all moments α^n and β^n, and all normally ordered averages are initially zero for the case studied here. By extension, all moments involving the initial values $x_1(0)$, $x_2(0)$, $p_1(0)$, and $p_2(0)$ are zero.

The formal solution [Eqs. (6.12)] yields an approximate solution, valid for short interaction times and large pump amplitude, when the stochastic variables are expanded in a perturbation series in the reciprocal of the pump amplitude:

$$\theta=\sum_{n=0}^{\infty}\alpha_0^{-n}\,\theta^{(n)}\ . \tag{6.14}$$

By substituting the expansion Eq. (6.14) into the formal solution Eqs. (6.12) and equating equal powers of α_0^{-n}, we obtain an approximate solution to the set of SDE's:

(i) to zeroth order,

$$z_1^{(0)}(u) = \frac{1}{\sqrt{2}} \int_0^u e^{-x}\, dV \ , \tag{6.15a}$$

$$z_2^{(0)}(u) = -\frac{i}{\sqrt{2}} \int_0^u e^{x}\, dW \ , \tag{6.15b}$$

$$p_1^{(0)}(u) = p_2^{(0)}(u) = 0 \ ; \tag{6.15c}$$

(ii) to first order,

$$z_1^{(1)}(u) = z_2^{(1)}(u) = 0 \ , \tag{6.16a}$$

$$p_1^{(1)}(u) = \frac{1}{2} \int_0^u [z_2^{(0)2}(x) e^{-2x} - z_1^{(0)2}(x) e^{2x}]\, dx \ , \tag{6.16b}$$

$$p_2^{(1)}(u) = -\int_0^u z_1^{(0)}(x) z_2^{(0)}(x)\, dx \ ; \tag{6.16c}$$

and (iii) to second order,

$$z_1^{(2)}(u) = \int_0^u [z_1^{(0)}(x) p_1^{(1)}(x) + z_2^{(0)}(x) p_2^{(1)}(x) e^{-2x}]\, dx$$

$$+ \frac{1}{2\sqrt{2}} \int_0^u e^{-x} [p_1^{(1)}(x)\, dV + i p_2^{(1)}(x)\, dW] \ , \tag{6.17a}$$

$$z_2^{(2)}(u) = \int_0^u [z_1^{(0)}(x) p_2^{(1)}(x) e^{2x} - z_2^{(0)}(x) p_1^{(1)}(x)]\, dx$$

$$- \frac{i}{2\sqrt{2}} \int_0^u e^{x} [p_1^{(1)}(x)\, dW + i p_2^{(1)}(x)\, dV] \ , \tag{6.17b}$$

$$p_1^{(2)}(u) = p_2^{(2)}(u) = 0 \ . \tag{6.17c}$$

The SDE's corresponding to the zeroth order solutions $z_1^{(0)}(u)$ and $z_2^{(0)}(u)$ are

$$dz_1^{(0)} = \frac{1}{\sqrt{2}} e^{-u}\, dV \ , \tag{6.18a}$$

$$dz_2^{(0)} = -\frac{i}{\sqrt{2}} e^{u}\, dW \ . \tag{6.18b}$$

We define the new variables

$$A_1(u) = z_1^{(0)}(u)\, e^{u} \ , \tag{6.19a}$$

$$A_2(u) = i z_2^{(0)}(u)\, e^{-u} \ , \tag{6.19b}$$

with the resulting SDE's

$$dA_1 = A_1\, du + \frac{1}{\sqrt{2}}\, dV \ , \tag{6.20a}$$

$$dA_2 = -A_2\, du + \frac{1}{\sqrt{2}}\, dW \ . \tag{6.20b}$$

Equations (6.20) describe two independent Ornstein-Uhlenbeck processes, each with zero mean. Thus, $z_1^{(0)}(u)$ and $z_2^{(0)}(u)$, apart from the exponential factors e^{u} and e^{-u}, respectively, are themselves Ornstein-Uhlenbeck processes. They are Gaussian variables; all higher-order moments can be expressed in terms of second order moments. With this in mind, we can formulate a pair of rules to guide us through the remaining calculations: (i) a rule for quadratic moments,

$$\langle z_1^{(0)}(u) z_1^{(0)}(w) \rangle_{\mathrm{av}} = \frac{1}{4}(1 - e^{-2w}) \quad u \geq w \ , \tag{6.21a}$$

$$\langle z_2^{(0)}(u) z_2^{(0)}(w) \rangle_{\mathrm{av}} = -\frac{1}{4}(e^{2w} - 1) \quad u \geq w \ , \tag{6.21b}$$

$$\langle z_1^{(0)}(u) z_2^{(0)}(w) \rangle_{\mathrm{av}} = 0 \ , \tag{6.21c}$$

and (ii) a rule for quartic moments,

$$<z_1^{(0)}(u)z_1^{(0)}(v)z_1^{(0)}(w)z_1^{(0)}(z)>_{av} = <z_1^{(0)}(u)z_1^{(0)}(v)>_{av}<z_1^{(0)}(w)z_1^{(0)}(z)>_{av}$$
$$+ <z_1^{(0)}(u)z_1^{(0)}(w)>_{av}<z_1^{(0)}(v)z_1^{(0)}(z)>_{av}$$
$$+ <z_1^{(0)}(u)z_1^{(0)}(z)>_{av}<z_1^{(0)}(v)z_1^{(0)}(w)>_{av},$$

(6.22a)

$$<z_2^{(0)}(u)z_2^{(0)}(v)z_2^{(0)}(w)z_2^{(0)}(z)>_{av} = <z_2^{(0)}(u)z_2^{(0)}(v)>_{av}<z_2^{(0)}(w)z_2^{(0)}(z)>_{av}$$
$$+ <z_2^{(0)}(u)z_2^{(0)}(w)>_{av}<z_2^{(0)}(v)z_2^{(0)}(z)>_{av}$$
$$+ <z_2^{(0)}(u)z_2^{(0)}(z)>_{av}<z_2^{(0)}(v)z_2^{(0)}(w)>_{av},$$

(6.22b)

where $< >_{av}$ denotes an average in the positive-P representation.

The squeezing in the signal mode is easily calculated by the repeated application of (i) and (ii). The signal-mode quadrature-phase amplitudes are defined by Eqs. (2.11), and the pump-mode quadrature-phase amplitudes are defined by Eq. (3.5), or more explicitly by

$$\hat{P}_1 = \frac{1}{2}(\hat{a}_p + \hat{a}_p^\dagger), \qquad \hat{P}_2 = -\frac{i}{2}(\hat{a}_p - \hat{a}_p^\dagger) \tag{6.23}$$

The expectation values of the signal-mode quadrature-phase amplitudes $\hat{x}_1$ and $\hat{x}_2$ are zero when the signal is initially vacuum, as shown by Eq. (3.11). We then find that the uncertainties in $\hat{x}_1$ and $\hat{x}_2$ are

$$<\Delta\hat{x}_1^2> = \frac{1}{4} + <x_1^2(u)>_{av} = \frac{1}{4} + <z_1^2(u)>_{av} e^{2u}, \tag{6.24a}$$

$$<\Delta\hat{x}_2^2> = \frac{1}{4} + <x_2^2(u)>_{\rm av} = \frac{1}{4} + <z_2^2(u)>_{\rm av} e^{-2u} \ . \qquad (6.24b)$$

We see from Eqs (6.24) that x_1 and x_2 are the c-number equivalents of the quadrature-phase amplitudes $\hat{x}_1$ and $\hat{x}_2$, respectively. To second order in α_0^{-1}, the uncertainties are

$$<\Delta\hat{x}_1^2> = \frac{1}{4} + <z_1^{(0)2}(u)>_{\rm av} e^{2u} + \frac{1}{\alpha_0^2} <2z_1^{(0)}(u) z_1^{(2)}(u)>_{\rm av} e^{2u} \ , \qquad (6.25a)$$

$$<\Delta\hat{x}_2^2> = \frac{1}{4} + <z_2^{(0)2}(u)>_{\rm av} e^{-2u} + \frac{1}{\alpha_0^2} <2z_2^{(0)}(u) z_2^{(2)}(u)>_{\rm av} e^{-2u} \ . \qquad (6.25b)$$

Application of (i) yields the ideal squeezing:

$$<\Delta\hat{x}_1^2>_{\rm ideal} = \frac{1}{4} + <z_1^{(0)2}(u)>_{\rm av} e^{2u} = \frac{1}{4} e^{2u} \ , \qquad (6.26a)$$

$$<\Delta\hat{x}_2^2>_{\rm ideal} = \frac{1}{4} + <z_2^{(0)2}(u)>_{\rm av} e^{-2u} = \frac{1}{4} e^{-2u} \ . \qquad (6.26b)$$

Repeated applications of (i) and (ii) yield the quadrature variances correct to semiclassical order $\alpha_0^{-2} = \bar{N}^{-1}$:

$$<\Delta\hat{x}_1^2> = \frac{1}{4} e^{2u} + \frac{1}{8\bar{N}} \Big[u^2 e^{2u} + u(e^{2u}+1)$$

$$- (3\sinh^2 u + 2)\sinh u \, e^u - \sinh^2 u \Big] \ , \qquad (6.27a)$$

$$<\Delta \hat{x}_2^2> = \frac{1}{4} e^{-2u} + \frac{1}{8\bar{N}} \Big[u^2 e^{-2u} - u(e^{-2u} + 1)$$

$$+ (3\sinh^2 u + 2)\sinh u \; e^{-u} - \sinh^2 u \Big] \; , \qquad (6.27b)$$

which are exactly the results obtained by Hillery and Zubairy.[7] Equation (6.27a) also agrees to $O(\bar{N}^{-1})$, with the result calculated via the seminumerical method [Eq. (5.8)]. The variance of the squeezed quadrature, including the dominant correction for at best moderate squeezing only, is

$$<\Delta \hat{x}_2^2> = \frac{1}{4} e^{-2u} + \frac{1}{64\bar{N}} e^{2u} \; , \qquad (6.28)$$

which agrees with Eq. (2.18), validating our heuristic picture of the effects of pump fluctuations on squeezing.

B. The two-mode PA

The analysis of the two-mode PA is similar to that of the one-mode PA. The equation of motion is again von Neumann's equation in the interaction representation, Eq. (6.1), with the two-mode interaction Hamiltonian $\hat{H}_{\rm int}(t)$ given by Eq. (1.4). Initially, we assume that the signal modes are in vacuum states, and that the pump mode is in a coherent state:

$$\hat{\rho}(0) = \hat{\rho}_I(0) = |0;0;\alpha_0><0;0;\alpha_0| \; . \qquad (6.29)$$

By substituting the two-mode versions of Eqs. (6.3) and (6.4) into Eq. (6.1) and

integrating by parts, we find the Fokker-Planck equation for the two-mode PA:

$$\frac{\partial P}{\partial \tau} = \Bigg[-\beta_2\alpha_p \frac{\partial}{\partial \alpha_1} - \alpha_2\beta_p \frac{\partial}{\partial \beta_1} - \beta_1\alpha_p \frac{\partial}{\partial \alpha_2} - \alpha_1\beta_p \frac{\partial}{\partial \beta_2}$$

$$+ \alpha_1\alpha_2 \frac{\partial}{\partial \alpha_p} + \beta_1\beta_2 \frac{\partial}{\partial \beta_p} + \alpha_p \frac{\partial^2}{\partial \alpha_1 \partial \alpha_2} + \beta_p \frac{\partial^2}{\partial \beta_1 \partial \beta_2} \Bigg] P \ , \tag{6.30}$$

where $\tau = \kappa t$ and $P \equiv P(\alpha_1, \alpha_2, \alpha_p, \beta_1, \beta_2, \beta_p, \tau)$.

Proceeding as in the one-mode case, we can derive a set of Ito SDE's from the Fokker-Planck equation, Eq. (6.30):

$$d\alpha_1 = \beta_2\alpha_p \, d\tau + \sqrt{\alpha_p} \, dW_1 \ , \tag{6.31a}$$

$$d\alpha_2 = \beta_1\alpha_p \, d\tau + \sqrt{\alpha_p} \, dW_1^* \ , \tag{6.31b}$$

$$d\beta_1 = \alpha_2\beta_p \, d\tau + \sqrt{\beta_p} \, dW_2 \ , \tag{6.31c}$$

$$d\beta_2 = \alpha_1\beta_p \, d\tau + \sqrt{\beta_p} \, dW_2^* \ , \tag{6.31d}$$

$$d\alpha_p = -\alpha_1\alpha_2 \, d\tau \ , \tag{6.31e}$$

$$d\beta_p = -\beta_1\beta_2 \, d\tau \ . \tag{6.31f}$$

The *complex* Wiener increments dW_1 and dW_2 are defined by

$$dW_1 = \frac{dW_{1x} + i dW_{1y}}{\sqrt{2}} \ , \tag{6.32a}$$

$$dW_2 = \frac{dW_{2x} + i dW_{2y}}{\sqrt{2}} \ , \tag{6.32b}$$

where dW_{1x}, dW_{1y}, dW_{2x}, and dW_{2y} are independent, real Wiener increments, and,

in the Ito calculus,

$$dW_{1s}^2 = dW_{2s}^2 = d\tau \qquad s = x, y \ . \tag{6.32c}$$

The complex pump amplitudes α_p and β_p are again assumed to consist of a large, real mean value α_0 plus small fluctuations, as in Eqs. (6.7). We define the new variables X_1, Y_1, X_2, Y_2, P_1, and P_2 by

$$X_1 = \frac{1}{2}(\alpha_1 + \beta_2) \ , \quad X_2 = -\frac{i}{2}(\alpha_1 - \beta_2) \ , \tag{6.33a}$$

$$Y_1 = \frac{1}{2}(\alpha_2 + \beta_1) \ , \quad Y_2 = -\frac{i}{2}(\alpha_2 - \beta_1) \ , \tag{6.33b}$$

$$P_1 = \frac{1}{2}(\Delta\alpha + \Delta\beta) \ , \ P_2 = -\frac{i}{2}(\Delta\alpha - \Delta\beta) \ . \tag{6.33c}$$

It is convenient to change variables one more time:

$$U_1 = X_1 e^{-u} \ , \quad U_2 = X_2 e^{u} \ , \tag{6.34a}$$

$$Z_1 = Y_1 e^{-u} \ , \quad Z_2 = Y_2 e^{u} \ , \tag{6.34b}$$

where $u = \alpha_0 \tau$. The resulting SDE's are

$$dU_1 = \frac{1}{\alpha_0}(U_1 P_1 + U_2 P_2 e^{-2u})\, du + \frac{1}{2} e^{-u} \left[\left(1 + \frac{P_1 + iP_2}{\alpha_0} \right)^{1/2} dV_1 \right.$$

$$\left. + \left(1 + \frac{P_1 - iP_2}{\alpha_0} \right)^{1/2} dV_2^* \right] , \tag{6.35a}$$

$$dZ_1 = \frac{1}{\alpha_0}(Z_1 P_1 + Z_2 P_2 e^{-2u})\,du + \frac{1}{2}e^{-u}\left[\left(1+\frac{P_1+iP_2}{\alpha_0}\right)^{1/2} dV_1^* + \left(1+\frac{P_1-iP_2}{\alpha_0}\right)^{1/2} dV_2\right], \tag{6.35b}$$

$$dU_2 = \frac{1}{\alpha_0}(U_1 P_2 e^{2u} - U_2 P_1)\,du - \frac{i}{2}e^{u}\left[\left(1+\frac{P_1+iP_2}{\alpha_0}\right)^{1/2} dV_1 - \left(1+\frac{P_1-iP_2}{\alpha_0}\right)^{1/2} dV_2^*\right], \tag{6.35c}$$

$$dZ_2 = \frac{1}{\alpha_0}(Z_1 P_2 e^{2u} - Z_2 P_1)\,du - \frac{i}{2}e^{u}\left[\left(1+\frac{P_1+iP_2}{\alpha_0}\right)^{1/2} dV_1^* - \left(1+\frac{P_1-iP_2}{\alpha_0}\right)^{1/2} dV_2\right], \tag{6.35d}$$

$$dP_1 = \frac{1}{\alpha_0}(U_2 Z_2 e^{-2u} - U_1 Z_1 e^{2u})\,du\,, \tag{6.35e}$$

$$dP_2 = -\frac{1}{\alpha_0}(U_1 Z_2 + U_2 Z_1)\,du\,, \tag{6.35f}$$

where $dV_1 = \sqrt{\alpha_0}\,dW_1$ and $dV_2 = \sqrt{\alpha_0}\,dW_2$.

We can obtain an approximate solution to Eqs. (6.35), just as we did in the one-mode case; we formally integrate Eqs. (6.35), expand the square roots, and substitute the expansion Eq. (6.14) . We have found the approximate solution up to second order in α_0^{-1}: (i) to zeroth order,

$$U_1^{(0)}(u) = \frac{1}{2}\int_0^u e^{-x}(dS_1 + i\,dS_2)\,, \tag{6.36a}$$

$$Z_1^{(0)}(u) = U_1^{(0)*}(u)\,, \tag{6.36b}$$

$$U_2^{(0)}(u) = \frac{1}{2}\int_0^u e^x (dS_3 - idS_4)\,, \tag{6.36c}$$

$$Z_2^{(0)}(u) = -U_2^{(0)*}(u)\,, \tag{6.36d}$$

$$P_1^{(0)}(u) = P_2^{(0)}(u) = 0\,; \tag{6.36e}$$

(ii) to first order,

$$U_1^{(1)}(u) = Z_1^{(1)}(u) = 0\,, \tag{6.37a}$$

$$U_2^{(1)}(u) = Z_2^{(1)}(u) = 0\,, \tag{6.37b}$$

$$P_1^{(1)}(u) = \int_0^u [U_2^{(0)}(x) Z_2^{(0)}(x) e^{-2x} - U_1^{(0)}(x) Z_1^{(0)}(x) e^{2x}]\, dx\,, \tag{6.37c}$$

$$P_2^{(1)}(u) = -\int_0^u [U_1^{(0)}(x) Z_2^{(0)}(x) + U_2^{(0)}(x) Z_1^{(0)}(x)]\, dx\,; \tag{6.37d}$$

and (iii) to second order,

$$U_1^{(2)}(u) = \int_0^u [U_1^{(0)}(x) P_1^{(1)}(x) + U_2^{(0)}(x) P_2^{(1)}(x) e^{-2x}]\, dx$$

$$+ \frac{1}{4}\int_0^u e^{-x} [P_1^{(1)}(x)(dS_1 + idS_2) - P_2^{(1)}(x)(dS_3 - idS_4)]\,, \tag{6.38a}$$

$$Z_1^{(2)}(u) = \int_0^u [Z_1^{(0)}(x) P_1^{(1)}(x) + Z_2^{(0)} P_2^{(1)}(x) e^{-2x}]\, dx$$

$$+ \frac{1}{4}\int_0^u e^{-x} [P_1^{(1)}(x)(dS_1 - idS_2) + P_2^{(1)}(x)(dS_3 + idS_4)]\,, \tag{6.38b}$$

$$U_2^{(2)}(u)=\int_0^u [U_1^{(0)}(x)P_2^{(1)}(x)e^{2x}-U_2^{(0)}(x)P_1^{(1)}(x)]\,dx$$

$$+\frac{1}{4}\int_0^u e^x [P_1^{(1)}(x)(dS_3-idS_4)+P_2^{(1)}(x)(dS_1+idS_2)]\,, \tag{6.38c}$$

$$Z_2^{(2)}(u)=\int_0^u [Z_1^{(0)}(x)P_2^{(1)}(x)e^{2x}-Z_2^{(0)}(x)P_1^{(1)}(x)]\,dx$$

$$-\frac{1}{4}\int_0^u e^x [P_1^{(1)}(x)(dS_3+idS_4)-P_2^{(1)}(x)(dS_1-idS_2)]\,, \tag{6.38d}$$

$$P_1^{(2)}(u)=P_2^{(2)}(u)=0\,. \tag{6.38e}$$

Notice that in Eqs. (6.36) and Eqs. (6.38) we have replaced the complex noise increments by the real Wiener increments dS_1, dS_2, dS_3, and dS_4, where, using Eqs. (6.32),

$$dS_1=\frac{dW_{1x}+dW_{2x}}{\sqrt{2}}\,,\quad dS_2=\frac{dW_{1y}-dW_{2y}}{\sqrt{2}}\,, \tag{6.39a}$$

$$dS_3=\frac{dW_{1y}+dW_{2y}}{\sqrt{2}}\,,\quad dS_4=\frac{dW_{1x}-dW_{2x}}{\sqrt{2}}\,. \tag{6.39b}$$

Also note that, as in the one-mode case, we have dropped all contributions arising from the initial conditions.

By comparing Eqs. (6.36) with Eqs. (6.15), we see that the zeroth-order solutions $U_1^{(0)}(u)$, $Z_1^{(0)}(u)$, $U_2^{(0)}(u)$, and $Z_2^{(0)}(u)$ have real and imaginary parts that are Gaussian variables. Let

$$U_1^{(0)}(u)=Z_1^{(0)*}(u)=Q_1(u)+iQ_2(u)\,, \tag{6.40a}$$

$$U_2^{(0)}(u)=-Z_2^{(0)*}(u)=Q_3(u)+iQ_4(u)\,, \tag{6.40b}$$

where $Q_1(u)$, $Q_2(u)$, $Q_3(u)$, and $Q_4(u)$ are *independent* Gaussian variables with zero

mean. We can use Eqs. (6.40) to generalize the one-mode rules [Eqs. (6.21) and Eqs. (6.22)] to the two-mode case: (i) a rule for quadratic moments,

$$<Q_1(u)Q_1(w)>_{\rm av}=<Q_2(u)Q_2(w)>_{\rm av}=\frac{1}{8}(1-e^{-2w}) \quad u\geq w\ , \tag{6.41a}$$

$$<Q_3(u)Q_3(w)>_{\rm av}=<Q_4(u)Q_4(w)>_{\rm av}=\frac{1}{8}(e^{2w}-1) \quad u\geq w\ , \tag{6.41b}$$

$$<Q_i(u)Q_j(u)>_{\rm av}=0 \qquad i\neq j\ , \tag{6.41c}$$

and (ii) a rule for quartic moments,

$$\begin{aligned}<Q_1(u)Q_1(v)Q_1(w)Q_1(z)>_{\rm av}=&<Q_1(u)Q_1(v)>_{\rm av}<Q_1(w)Q_1(z)>_{\rm av}\\&+<Q_1(u)Q_1(w)>_{\rm av}<Q_1(v)Q_1(z)>_{\rm av}\\&+<Q_1(u)Q_1(z)>_{\rm av}<Q_1(v)Q_1(w)>_{\rm av}\ ,\end{aligned} \tag{6.42a}$$

$$<Q_2(u)Q_2(v)Q_2(w)Q_2(z)>_{\rm av}=<Q_1(u)Q_1(v)Q_1(w)Q_1(z)>_{\rm av}\ , \tag{6.42b}$$

$$\begin{aligned}<Q_3(u)Q_3(v)Q_3(w)Q_3(z)>_{\rm av}=&<Q_3(u)Q_3(v)>_{\rm av}<Q_3(w)Q_3(z)>_{\rm av}\\&+<Q_3(u)Q_3(w)>_{\rm av}<Q_3(v)Q_3(z)>_{\rm av}\\&+<Q_3(u)Q_3(z)>_{\rm av}<Q_3(v)Q_3(w)>_{\rm av}\ ,\end{aligned} \tag{6.42c}$$

$$<Q_4(u)Q_4(v)Q_4(w)Q_4(z)>_{\rm av}=<Q_3(u)Q_3(v)Q_3(w)Q_3(z)>_{\rm av}\ . \tag{6.42d}$$

We can calculate the two-mode squeezing by repeated application of (i) and (ii). The two-mode quadrature-phase amplitudes are defined by

$$\hat{X}_1=\frac{1}{2}(\hat{a}_1+\hat{a}_2^{\dagger})\ , \qquad \hat{X}_2=-\frac{i}{2}(\hat{a}_1-\hat{a}_2^{\dagger})\ , \tag{6.43a}$$

$$\hat{P}_1 = \frac{1}{2}(\hat{a}_p + \hat{a}_p^\dagger), \qquad \hat{P}_2 = -\frac{i}{2}(\hat{a}_p - \hat{a}_p^\dagger). \tag{6.43b}$$

The two-mode signal quadrature-phase amplitudes are *not* Hermitian operators. In terms of the stochastic c-numbers, the mean-square uncertainties in the two-mode quadrature-phase amplitudes are given by

$$<|\Delta\hat{X}_1|^2> = \frac{1}{4} + <X_1 Y_1>_{av} - <X_1>_{av}<Y_1>_{av}, \tag{6.44a}$$

$$<|\Delta\hat{X}_2|^2> = \frac{1}{4} + <X_2 Y_2>_{av} - <X_2>_{av}<Y_2>_{av}. \tag{6.44b}$$

From Eq. (1.13) and Eqs. (6.44), we see that X_1 and Y_1 are the c-number equivalents of the operators $\hat{X}_1$ and its Hermitian conjugate $\hat{X}_1^\dagger$, respectively, and X_2 and Y_2 are the c-number equivalents of the operators $\hat{X}_2$ and its Hermitian conjugate $\hat{X}_2^\dagger$. Substituting Eq. (6.14), Eqs. (6.36), Eqs. (6.37), and Eqs. (6.38) into Eqs. (6.44), we have to second order in α_0^{-1}

$$<|\Delta\hat{X}_1|^2> = \frac{1}{4} + <U_1^{(0)}(u) Z_1^{(0)}(u)>_{av} e^{2u}$$

$$+ <U_1^{(0)}(u) Z_1^{(2)}(u) + U_1^{(2)}(u) Z_1^{(0)}(u)>_{av} e^{2u}, \tag{6.45a}$$

$$<|\Delta\hat{X}_2|^2> = \frac{1}{4} + <U_2^{(0)}(u) Z_2^{(0)}(u)>_{av} e^{-2u}$$

$$+ \frac{1}{\alpha_0^2} <U_2^{(0)}(u) Z_2^{(2)}(u) + U_2^{(2)}(u) Z_2^{(0)}(u)>_{av} e^{-2u}, \tag{6.45b}$$

since the expectation values of X_1, Y_1, X_2, and Y_2 are zero.

Repeated application of the two-mode rules (i) and (ii) yields the mean-squared uncertainties correct to semiclassical order $\alpha_0^{-2}=\bar{N}^{-1}$:

$$<|\Delta\hat{X}_1|^2>=\frac{1}{4}e^{2u}+\frac{1}{8\bar{N}}\left[u^2e^{2u}+u(\frac{3}{2}e^{2u}+1)\right.$$

$$\left.-(4\sinh^2u+\frac{5}{2})\sinh u\ e^{u}-\frac{3}{2}\sinh^2u\right], \tag{6.46a}$$

$$<|\Delta\hat{X}_2|^2>=\frac{1}{4}e^{-2u}+\frac{1}{8\bar{N}}\left[u^2e^{-2u}-u(\frac{3}{2}e^{-2u}+1)\right.$$

$$\left.+(4\sinh^2u+\frac{5}{2})\sinh u\ e^{-u}-\frac{3}{2}\sinh^2u\right], \tag{6.46b}$$

which is slightly different from the one-mode result, Eqs. (6.27), and agrees to $O(\bar{N}^{-1})$ with the result calculated via the seminumerical method [Eq. (5.9)]. When the dominant correction only is kept, the uncertainty in the squeezed quadrature is

$$<|\Delta\hat{X}_2|^2>\cong\frac{1}{4}e^{-2u}+\frac{1}{64\bar{N}}e^{2u}, \tag{6.47}$$

which is the same as the dominant correction found for the one-mode case, Eq.(6.28), in contrast to the result obtained by Scharf and Walls.[8]

VII. CONCLUSION

We have calculated, for the one- and two-mode PA, the explicit corrections for squeezing to order $1/\bar{N}^2$, that are due to a quantum pump in a coherent state with an average photon number $\bar{N}$. We found that the pump's phase noise is responsible for the dominant contribution to the limitations on squeezing for any number of signal

modes. We also briefly discuss under what conditions traveling-wave calculations can be treated by Hamiltonian methods in the most direct way. This was done by discretizing the continuum problem. Finally, the limitation to squeezing in the discrete-mode calculations that we performed were shown to be insensitive to the details of our discretization process.

ACKNOWLEDGEMENTS

The authors would like to thank Carlton Caves for his useful suggestions. Thank you, Carl. This work was supported in part by Caltech's Program in Advanced Technologies, sponsored by Aerojet General, General Motors, and TRW, and by the Caltech President's Fund.

REFERENCES

[1] A. Yariv, *Quantum Electronics* (Wiley, New York, 1975), Chap. 17.

[2] R. G. Smith, in *Laser Handbook,* F. T. Arrechi and E. O. Shulz-Dubois, eds. (North Holland, Amsterdam, 1972), Vol. I, p. 837.

[3] H. Takahashi, Adv. Commun. Syst. **1**, 227 (1965), especially Section VI.

[4] E. Y. C. Lu, Lett. Nuovo Cimento **3**, 585 (1972).

[5] D. Stoler, Phys. Rev. Lett. **33**, 1397 (1974).

[6] K. Wódkiewicz and M. S. Zubairy, Phys. Rev. A **27**, 2003 (1983).

[7] M. Hillery and M. S. Zubairy, Phys. Rev. A **29**, 1275 (1984).

[8] G. Scharf and D. F. Walls, Opt. Commun. **50**, 245 (1984).

[9] M. Hillery and M. S. Zubairy, Phys. Rev. A **26**, 451 (1982).

[10] M. S. Zubairy, private communication.

[11] G. Scharf, Annals of Physics **83**, 71 (1974).

[12] G. Scharf and R. Weiss, Helvetica Physica Acta **47**, 505 (1974).

[13] G. Scharf, Helvetica Physica Acta **48**, 329 (1975); **50**, 253 (1977).

[14] C. M. Caves and B. L. Schumaker, Phys. Rev. A **31**, 3068 (1985).

[15] C. M. Caves and D. D. Crouch, J. Opt. Soc. Am. B **4**, 1535 (1987).

[16] P. D. Drummond and C. W. Gardiner, J. Phys. A **13**, 2353 (1980).

[17] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1983); L. Arnold, *Stochastic Differential Equations* (Wiley, New York, 1973).

[18] J. Tucker and D. F. Walls, Phys. Rev. **178**, 2036 (1969).

[19] A. Lane, P. Tombesi, H. J. Carmichael, and D. F. Walls, Opt. Commun **48**, 155 (1983).

[20] C. M. Caves and B. L. Schumaker, in *Quantum Optics IV,* J. D. Harvey and D. F.

Walls, eds., Springer Proceedings in Physics (Springer-Verlag, New York, 1986).

[21] H. J. Carmichael and M. Wolinsky, in *Quantum Optics IV*, J. D. Harvey and D. F. Walls, eds., Springer Proceedings in Physics (Springer-Verlag, New York, 1986).

FIGURE CAPTIONS

FIG. 1. The effect of pump phase fluctuations on squeezing. The ellipse with solid lines represents ideal squeezing, in which the pump has a well-defined phase. Phase fluctuations in the pump cause the orientation of the ellipse to fluctuate about $\phi_p = 0$, with the characteristic angle, $\Delta\phi = 1/(4\bar{N}^{1/2})$, feeding noise from the amplified quadrature into the squeezed quadrature.

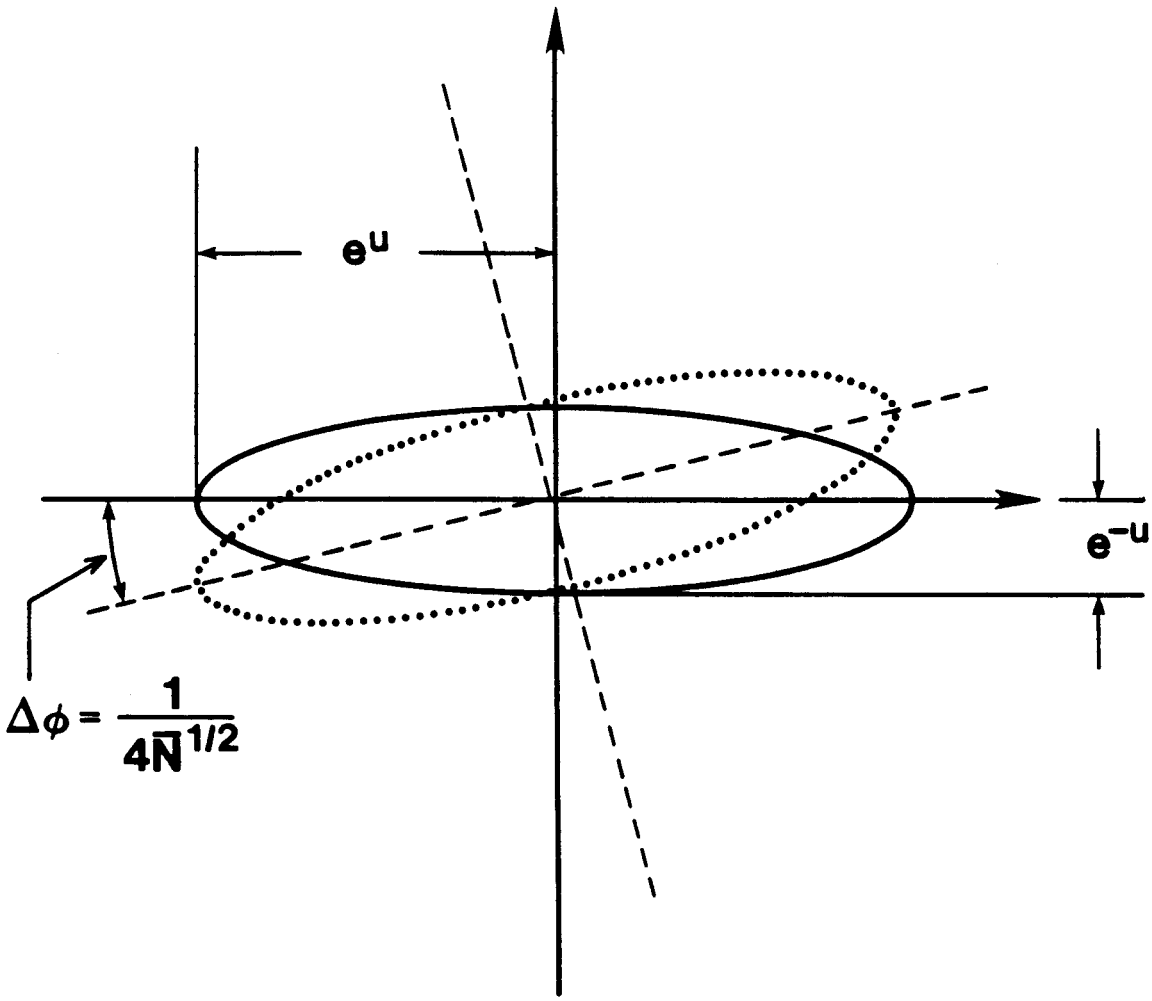

Figure 1

PART TWO

CHAPTER 4

Quantum rules: An Effect can have more than one Operation

by Samuel L. Braunstein and Carlton M. Caves

Accepted for publication by Foundations of Physics Letters.

ABSTRACT

We review the formalism of Effects and Operations in order to demonstrate that this formalism is equivalent to the standard rules of quantum mechanics. D'Espagnat[5] has studied an example where he finds a discrepancy between an analysis based on the standard quantum rule for calculating probabilities and an analysis that uses a particular Operation taken from the work of Barchielli and his collaborators.[4] We use the formalism of Effects and Operations to explore and explain this discrepancy.

1. INTRODUCTION

Measurements in quantum mechanics are usually described in the following way: the system to be measured is coupled to a "measuring apparatus"; after the interaction, one obtains information about the system by observing some property of the apparatus. We shall refer to such a description as a "measurement model."

Quantum-mechanical measurements can also be described by a formalism that uses mathematical objects called Effects and Operations.[1–3] The formalism of Effects and Operations is equivalent to a description in terms of a measurement model. Recently, however, there has been some question about the usage of this formalism.[4,5] The purpose of this brief note is to clear up this question.

The formalism of Effects and Operations can be obtained by applying the standard rules of quantum mechanics to the observation of the apparatus in a measurement model. What the formalism does is to provide a very convenient, compact notation for describing a measurement directly in terms of the system state (density operator). The properties of a particular model—i.e., the type of apparatus and its quantum state and the type of interaction with the system—are incorporated in the Effects and Operations.

A set of Effects generates from the system state the statistics of the possible outcomes. Given a particular result, an Operation maps the system state before the measurement to a new (unnormalized) system state after the measurement. Thus, one obtains a complete quantum-mechanical description of the measurement. Different models can give rise to the same measurement statistics, yet lead to different system states after the measurement. Thus, one says that an Effect can have more than one Operation. The difference between two such models shows up in the statistics of a second measurement.

Barchielli, Lanz, and Prosperi[4] have used Effects and Operations to develop an elegant formal description of a continuous measurement of position. To generate the measurement statistics, they use a simple set of Effects. More importantly, they adopt the simplest Operations that are consistent with their Effects. In particular, they use Operations, called pure Operations, that map pure system states to pure system states, though even within the class of pure Operations their choice is special.

Barchielli *et al.*[4] do not give a measurement model that realizes their choice of Effects and Operations (the reader can find such a model in Ref. 6), although there is a theorem that any consistent set of Effects and Operations can be realized.[1] Even a careful reading of their paper might lead one to believe that they are proposing their particular choice of Operations as a fundamental constituent in a new theory of measurement. Only in their Appendix B do they make it clear that their choice of Operations is not unique; it corresponds to some measurement model that they do not specify.

In a recent paper, d'Espagnat[5] takes the pure Operation of Barchielli *et al.* as part of a new theory, which he dubs the "Milano theory." Seeking to compare this theory with standard quantum mechanics, he analyzes two consecutive measurements

of spin by two different methods. In the first method, he formulates a model for measurements of spin and applies to the model the standard quantum rule for calculating probabilities. In the second method he applies to two measurements of spin a pure Operation analogous to the pure Operation for position measurements of Barchielli *et al.* He finds that the two methods do not agree and concludes that the "Milano theory" is different from standard quantum theory. In view of the above discussion, however, the discrepancy he finds is not at all surprising. The pure Operation of the "Milano theory" corresponds to *some* model for measurements of spin, but it is not the measurement model that d'Espagnat has formulated. D'Espagnat knocks down a straw man—the "Milano theory". The formalism of Effects and Operations remains untouched.

In this note we explore and explain the discrepancy found by d'Espagnat. In Section 2, we review briefly the formalism of Effects and Operations, showing how it arises naturally from applying the standard quantum rules to a measurement model. In Section 3, we study d'Espagnat's model in detail and discuss how it is handled within the framework of Effects and Operations. In Section 4, we consider the pure Operation for a spin measurement (the "Milano theory"), and we formulate a measurement model, different from d'Espagnat's, that realizes this pure Operation.

2. REVIEW OF EFFECTS AND OPERATIONS

Consider a measurement model in which the system, initially with density operator $\hat{\rho}$, interacts with a measuring apparatus, initially with density operator $\hat{\rho}_{\mathcal{A}}$. The interaction leaves the system and the apparatus with joint density operator

$$\hat{\rho}_{\text{tot}} = \hat{U}\,\hat{\rho}\otimes\hat{\rho}_{\mathcal{A}}\hat{U}^{\dagger}\,, \tag{2.1}$$

where $\hat{U}$ is the joint evolution operator including the interaction.

After the interaction, one observes the apparatus in order to learn something about the system. Suppose, for example, that one measures an apparatus observable $\hat{B}$, which has eigenvalues B. Let $\hat{P}_B$ be the projector onto the space spanned by the eigenvectors with eigenvalue B. (An eigenvalue B can be degenerate, in which case $\hat{P}_B$ projects into a multidimensional subspace of the apparatus Hilbert space.) The standard quantum rule for calculating probabilities gives the probability to obtain result B:

$$P(B) = \mathrm{tr}(\hat{P}_B \hat{\rho}_{\mathrm{tot}}) = \mathrm{tr}_{\mathcal{S}}(\hat{F}_B \hat{\rho}) , \tag{2.2}$$

where

$$\hat{F}_B \equiv \mathrm{tr}_{\mathcal{A}}(\hat{U}^{\dagger} \hat{P}_B \hat{U} \hat{\rho}_{\mathcal{A}}) . \tag{2.3}$$

Here $\mathrm{tr}_{\mathcal{S}}$ and $\mathrm{tr}_{\mathcal{A}}$ denote traces over system and apparatus variables, respectively, and tr denotes a trace over both sets of variables. The system operators $\hat{F}_B$ are called Effects;[1–3] they provide a compact notation for generating measurement statistics from the system density operator.

Suppose now that the measurement yields the particular result B. The standard quantum rule for state reduction gives the system density operator $\hat{\rho}_B$ after the measurement: project the joint density operator into the subspace corresponding to eigenvalue B, trace out the apparatus, and normalize. The resulting state is

$$\hat{\rho}_B = \frac{\mathrm{tr}_{\mathcal{A}}(\hat{P}_B \hat{\rho}_{\mathrm{tot}} \hat{P}_B)}{P(B)} = \frac{\mathcal{F}_B \hat{\rho}}{P(B)} , \tag{2.4}$$

where $\mathcal{F}_B$ is a linear mapping on system density operators, defined by

$$\mathcal{F}_B\hat{\rho} \equiv \mathrm{tr}_{\mathcal{A}}(\hat{P}_B\hat{\rho}_{\mathrm{tot}}\hat{P}_B) = \mathrm{tr}_{\mathcal{A}}(\hat{P}_B\hat{U}\hat{\rho}\otimes\hat{\rho}_{\mathcal{A}}\hat{U}^{\dagger}) . \tag{2.5}$$

The map $\mathcal{F}_B$ is called an Operation;[1,3] it maps the system density operator before the measurement into the (unnormalized) system density operator after the measurement. An immediate consequence of the above is that

$$\mathrm{tr}_{\mathcal{S}}(\mathcal{F}_B\hat{\rho}) = \mathrm{tr}_{\mathcal{S}}(\hat{F}_B\hat{\rho}) = P(B) . \tag{2.6}$$

We emphasize again that Effects and Operations arise from applying the standard quantum rules to a measurement model—i.e., from applying projection operators to the measuring apparatus. Indeed, Operations generalize the usual state reduction to "indirect measurements," where a projection operator is applied to the apparatus rather than directly to the system.

An important special case arises when the apparatus is initially in a pure state $\hat{\rho}_{\mathcal{A}} = |\mathrm{Y}\rangle\langle\mathrm{Y}|$, and the projection operator $\hat{P}_B = |B\rangle\langle B|$ projects into a one-dimensional subspace of the apparatus Hilbert space. Then, defining a system operator $\hat{\mathrm{Y}}_B \equiv \langle B|\hat{U}|\mathrm{Y}\rangle$, one finds that $\hat{F}_B = \hat{\mathrm{Y}}_B^{\dagger}\hat{\mathrm{Y}}_B$ and $\mathcal{F}_B\hat{\rho} = \hat{\mathrm{Y}}_B\hat{\rho}\hat{\mathrm{Y}}_B^{\dagger}$. In this case, the Operation maps pure states to (unnormalized) pure states, and it is called a pure Operation.[3] Still more special is the case where $\hat{\mathrm{Y}}_B$ is Hermitian, so that $\mathcal{F}_B\hat{\rho} = \hat{F}_B^{1/2}\hat{\rho}\hat{F}_B^{1/2}$.

An Operation determines the corresponding Effect via Eq. (2.6), but the reverse is not true. Many Operations lead to the same Effect. Physically, we would say that many measurement models yield the same measurement statistics for a single measurement, even though their internal workings are different. The Effect $\hat{F}_B$ is consistent with the pure Operation $\hat{F}_B^{1/2}\hat{\rho}\hat{F}_B^{1/2}$, but many other Operations—pure and impure—produce the same Effect.

How could one distinguish between different Operations that lead to the same Effect? [Not by the statistics of a single measurement, because Effects determine those statistics.] One looks, instead, at the statistics of two or more successive measurements. Different Operations—i.e., different measurement models—lead to different post-measurement system states and thus, in general, to different statistics for a second measurement.

Barchielli, Lanz, and Prosperi[4] use the pure Operation $\hat{F}_B^{1/2}\hat{\rho}\hat{F}_B^{1/2}$ in their formal description of continuous position measurements. This choice corresponds to a particular measurement model,[6] which they do not specify. D'Espagnat[5] dubs this choice the "Milano theory" and seeks to compare it with standard quantum mechanics. He formulates a model for measurements of spin and shows that the "Milano theory" yields different statistics for two consecutive measurements of spin than he obtains when he applies to his model the standard quantum rule for calculating probabilities. The discrepancy he finds, while correct, has a simple explanation: The pure Operation $\hat{F}_B^{1/2}\hat{\rho}\hat{F}_B^{1/2}$ corresponds to a different measurement model than the model d'Espagnat has formulated.

We proceed in Section 3 to show how d'Espagnat's model is treated in terms of Effects and Operations.

3. D'ESPAGNAT'S MODEL

In this section, we consider the measurement model discussed by d'Espagnat,[5] with attention paid to its description in terms of Effects and Operations. The model also makes explicit the general formalism sketched in Section 2.

The system is a spin-½ particle, whose free-evolution is described by a Hamiltonian $\hat{H}_0$; the corresponding free evolution operator is $\hat{U}_0(t)=\exp(-i\hat{H}_0 t)$. The

measuring apparatus is a one-dimensional quantum system; its position x serves as a pointer that provides information about the z-component of the particle's spin. The apparatus is coupled to the particle strongly for a short time; we idealize the interaction as a δ-function in time. Thus, the total Hamiltonian is

$$\hat{H} = \hat{H}_0 + \alpha\,\delta(t - t_1)\,\hat{\sigma}_z \hat{p}\ , \tag{3.1}$$

where t_1 is the time when the apparatus is coupled to the particle, $\alpha > 0$ is a coupling constant, $\hat{\sigma}_z$ is the z-component of the particle's spin (in units of $\frac{1}{2}\hbar$), and $\hat{p}$ is the momentum conjugate to the pointer position $\hat{x}$. We assume that the apparatus has large enough mass that its free evolution can be neglected.

Let $|\phi_1\rangle$ and $|\phi_2\rangle$ be the eigenstates of $\hat{\sigma}_z$, corresponding to spin-up and spin-down, respectively:

$$\hat{\sigma}_z\,|\phi_j\rangle = (-1)^{j+1}\,|\phi_j\rangle\ , \quad j=1,2\ . \tag{3.2}$$

Suppose that at $t = 0$ the particle is in a pure state $|\psi_0\rangle$, and that the apparatus is prepared in a state with wave function $g(x)$. Then, just after the interaction, the joint state of the particle and apparatus becomes

$$\sum_j g(x - a_j)\,|\phi_j\rangle\langle\phi_j|\psi_1\rangle\ , \tag{3.3}$$

where $|\psi_1\rangle \equiv \hat{U}_0(t_1)\,|\psi_0\rangle$ is the particle's state just before the interaction at time t_1 and

$$a_j \equiv (-1)^{j+1}\,\alpha\ , \quad j=1,2\ . \tag{3.4}$$

The interaction displaces the pointer a distance α—to the right if the particle's z-spin is up, to the left if the particle's z-spin is down. The probability density to find the

pointer at position x is

$$P(x)=\sum_j |g(x-a_j)|^2 |\langle\phi_j|\psi_1\rangle|^2$$

$$= \psi_1 \left| \left[\sum_j |\phi_j\rangle |g(x-a_j)|^2 \langle\phi_j| \right] \right| \psi_1 \quad . \tag{3.5}$$

Suppose now that at time t_2 the particle interacts in the same way with a second, identical apparatus, whose position is labeled by y. Just after the second interaction, the joint state of the system and the two apparatuses is

$$\sum_{i,j} g(y-a_i)g(x-a_j)|\phi_i\rangle\langle\phi_i|\hat{U}_{21}|\phi_j\rangle\langle\phi_j|\psi_1\rangle \ , \tag{3.6}$$

where $\hat{U}_{21} \equiv \hat{U}_0(t_2-t_1)$. The joint probability density for the two-pointer positions is

$$P(x,y)=\sum_{i,j,k} |g(y-a_i)|^2 g^*(x-a_k)g(x-a_j)$$

$$\times \langle\psi_1|\phi_k\rangle\langle\phi_k|\hat{U}_{21}^{\dagger}|\phi_i\rangle\langle\phi_i|\hat{U}_{21}|\phi_j\rangle\langle\phi_j|\psi_1\rangle \ . \tag{3.7}$$

D'Espagnat considers the case where one asks only whether the pointer lies to the right of the origin or to the left of the origin. This corresponds to measuring an apparatus observable $\hat{P}_{(+)}-\hat{P}_{(-)}$, which has eigenvalues ± 1; the projection operators $\hat{P}_{(+)}$ and $\hat{P}_{(-)}$ are defined by

$$\hat{P}_I \equiv \int_I dx \ |x\rangle\langle x| \ , \tag{3.8}$$

where I can be either the interval $(+)\equiv(0,\infty)$ or the interval $(-)\equiv(-\infty,0)$. If the initial apparatus wave function is centered at the origin, and if the interaction displaces the pointer a distance somewhat larger than the width of the wave function, then knowing

on which side of the origin the pointer lies gives a good indication of whether the particle's z-spin is up or down.

The probability that the first pointer lies in the interval I is

$$P(I)=\int_I dx\, P(x)=\mathrm{tr}_{\mathcal{S}}(\hat{F}_I\,\hat{\rho}_1)\,. \tag{3.9}$$

Here, $\hat{\rho}_1 \equiv |\psi_1><\psi_1|$ is the particle's density operator just before the interaction at time t_1, and

$$\hat{F}_I \equiv \sum_j |\phi_j>\left[\int_I dx\ |g(x-a_j)|^2\right]<\phi_j| \tag{3.10}$$

is an Effect [cf. Eq. (3.5)]. Similarly, the joint probability that the first pointer lies in the interval I_1, and the second lies in the interval I_2 is

$$\begin{aligned} P(I_1,I_2)&=\int_{I_1} dx \int_{I_2} dy\, P(x,y) \\ &= \sum_{i,j,k}\left[\int_{I_2} dy\ |g(y-a_i)|^2\right]\left[\int_{I_1} dx\, g^*(x-a_k)g(x-a_j)\right] \\ &\qquad \times <\psi_1|\phi_k><\phi_k|\hat{U}_{21}^{\dagger}|\phi_i><\phi_i|\hat{U}_{21}|\phi_j><\phi_j|\psi_1>\,. \end{aligned} \tag{3.11}$$

The probabilities (3.9) and (3.11) come directly from the standard quantum rule for calculating probabilities; in particular, to obtain $P(I_1,I_2)$, there is no need to invoke state reduction after the first measurement.

How would the same two measurements be described, using Effects and Operations? Equation (3.9) gives the probability for the first measurement in terms of an Effect $\hat{F}_I$. If the first pointer is found to lie in the interval I, then the state of the particle just after the first measurement is

$$\hat{\rho}_I = \mathcal{F}_I \hat{\rho}_1 / P(I) , \tag{3.12}$$

where the Operation $\mathcal{F}_I$ is defined by

$$\mathcal{F}_I \hat{\rho}_1 = \sum_{j,k} \int_I dx \; g(x - a_j) g^*(x - a_k) |\phi_j\rangle \langle\phi_j| \hat{\rho}_1 |\phi_k\rangle \langle\phi_k| . \tag{3.13}$$

Notice that, in general, $\hat{\rho}_I$ is a mixed state. Physically, this is because one throws away information when one asks only on which side of the origin the pointer lies. Formally, it is because $\hat{P}_{(+)}$ and $\hat{P}_{(-)}$ are not one-dimensional projectors.

The second measurement is identical to the first. Thus, the conditional probability that the second pointer lies in the interval I_2, given that the first lay in I_1, is

$$P(I_2 | I_1) = \mathrm{tr}_{\mathcal{S}} (\hat{F}_{I_2} \hat{U}_{21} \hat{\rho}_{I_1} \hat{U}_{21}^{\dagger}) . \tag{3.14}$$

The joint probability for the two measurements can now be derived as

$$P(I_1, I_2) = P(I_2 | I_1) P(I_1) = \mathrm{tr}_{\mathcal{S}} (\hat{F}_{I_2} \hat{U}_{21} \mathcal{F}_{I_1} \hat{\rho}_1 \hat{U}_{21}^{\dagger}) . \tag{3.15}$$

One can easily see that Eq. (3.15) gives the same joint probability as Eq. (3.11). The quantum rule for calculating probabilities goes directly to this joint probability. The formalism of Effects and Operations proceeds more indirectly: a probability for the first measurement, followed by a state reduction described by an Operation, then a conditional probability for the second measurement, and, finally, the joint probability for the two measurements. Direct or indirect, the results are the same, because they are constructed to be the same.

After formulating and analyzing his model, d'Espagnat goes on to analyze two consecutive measurements of spin using the special pure Operation $\hat{F}_I^{1/2} \hat{\rho} \hat{F}_I^{1/2}$, which he takes from Barchielli *et al.* and which he calls the "Milano theory." We now

proceed in Section 4 to consider this pure Operation and to formulate a measurement model that corresponds to it.

4. THE PURE OPERATION AND A MODEL FOR IT

Consider now the special pure Operation $\mathcal{K}_I$ discussed in Section 2:

$$\mathcal{K}_I\hat{\rho}_1 \equiv \hat{F}_I^{1/2}\hat{\rho}_1\hat{F}_I^{1/2}$$

$$= \sum_{j,k}\left[\int_I dx\ |g(x-a_j)|^2\right]^{1/2}\left[\int_I dx\ |g(x-a_k)|^2\right]^{1/2}$$

$$\times|\phi_j><\phi_j|\hat{\rho}_1|\phi_k><\phi_k| \ . \tag{4.1}$$

This pure Operation clearly reproduces the Effect (3.10) for a single measurement in d'Espagnat's model—i.e., $\mathrm{tr}_{\mathcal{S}}(\mathcal{K}_I\hat{\rho}_1)=\mathrm{tr}_{\mathcal{S}}(\hat{F}_I\hat{\rho}_1)$; thus, it leads to the same measurement statistics for a single measurement. It should be compared and contrasted, however, with the actual Operation $\mathcal{F}_I$ [Eq. (3.13)] for the model.

Suppose that one uses the pure Operation $\mathcal{K}_I$ in place of the actual Operation $\mathcal{F}_I$ to analyze two consecutive measurements of spin. Using $\mathcal{K}_I$ to generate a new system state after the first measurement, one finds a joint probability for two measurements,

$$P'(I_1,I_2)=\mathrm{tr}_{\mathcal{S}}(\hat{F}_{I_2}\hat{U}_{21}\mathcal{K}_{I_1}\hat{\rho}_1\hat{U}_{21}^{\dagger})$$

$$= \sum_{i,j,k}\left[\int_{I_2} dx\ |g(x-a_i)|^2\right]$$

$$\times\left[\int_{I_1} dx\ |g(x-a_k)|^2\right]^{1/2}\left[\int_{I_1} dx\ |g(x-a_j)|^2\right]^{1/2}$$

$$\times<\psi_1|\phi_k><\phi_k|\hat{U}_{21}^{\dagger}|\phi_i><\phi_i|\hat{U}_{21}|\phi_j><\phi_j|\psi_1> \tag{4.2}$$

[cf. Eq. (3.15)]. As d'Espagnat points out, the joint probability (4.2) is different in general from the joint probability (3.11), which comes directly from applying the standard quantum rules to his model. D'Espagnat concludes that the "Milano theory" is different from standard quantum mechanics. Having learned about Effects and Operations, we can reach a different conclusion: The "Milano theory" is a straw man, because the pure Operation $\mathcal{K}_I$ is not meant to apply to all measurement models for spin; it is only one of many Operations that correspond to the Effect $\hat{F}_I$, and it is not the Operation that applies to d'Espagnat's model.

The two descriptions coincide when the initial apparatus state and the coupling constant are chosen so that the measurements are "ideal"—i.e., when $g(x)=0$ for $|x| > \alpha$. In this case, knowing the pointer lies to the right (left) of the origin tells one unambiguously that the particle's z-spin is up (down). The Effect $\hat{F}_I$ reduces to a projection operator—$\hat{F}_{(+)} = |\phi_1\rangle\langle\phi_1|$ and $\hat{F}_{(-)} = |\phi_2\rangle\langle\phi_2|$—and the corresponding Operations $\mathcal{F}_I = \mathcal{K}_I$ project onto $|\phi_1\rangle$ and $|\phi_2\rangle$.

More generally, it is interesting to inquire what measurement model does realize the pure Operation $\mathcal{K}_I$. Since the Operations $\mathcal{K}_{(+)}$ and $\mathcal{K}_{(-)}$ take a pure state $\hat{\rho}_1 = |\psi_1\rangle\langle\psi_1|$ to (unnormalized) pure states $\mathcal{K}_{(+)}\hat{\rho}_1$ and $\mathcal{K}_{(-)}\hat{\rho}_1$, it is convenient to represent their action in terms of state vectors instead of density operators:

$$|\psi_1\rangle \underset{\mathcal{K}_{(+)}}{\rightarrow} \Big(|\phi_1\rangle A \langle\phi_1| + |\phi_2\rangle B \langle\phi_2| \Big) |\psi_1\rangle \,, \tag{4.3a}$$

$$|\psi_1\rangle \underset{\mathcal{K}_{(-)}}{\rightarrow} \Big(|\phi_1\rangle (1-A^2)^{1/2} \langle\phi_1| + |\phi_2\rangle (1-B^2)^{1/2} \langle\phi_2| \Big) |\psi_1\rangle \,, \tag{4.3b}$$

where

$$A \equiv \left(\int_0^\infty dx \; |g(x-\alpha)|^2 \right)^{1/2} , \tag{4.4a}$$

$$B \equiv \left(\int_0^\infty dx \; |g(x+\alpha)|^2 \right)^{1/2} . \tag{4.4b}$$

The only condition on A and B is that they be between zero and one inclusive. (An "ideal" measurement would correspond to the case $A = 1$ and $B = 0$.)

The form (4.3) makes easy the search for a model yielding the Operation $\mathcal{A}_I$. Since the Operation is pure and since there are only two distinguished configurations for the apparatus, we can model the apparatus as another spin-½ system. We choose an interaction Hamiltonian

$$\hat{H}_{\text{int}} \equiv \hbar \beta \, \delta(t - t_1) \, \hat{\sigma}_z \hat{S}_y \; , \tag{4.5}$$

where $\beta > 0$ is a coupling constant, $\hat{\sigma}_z$ is again the z-component of the particle's spin, and $\hat{S}_y$ is the y-component of the apparatus's spin (in units of ½$\hbar$). The interaction causes the apparatus spin to precess about the y-axis—in the positive sense if the particle's z-spin is up, in the negative sense if the particle's z-spin is down.

The evolution operator corresponding to $\hat{H}_{\text{int}}$ is

$$\hat{U} = \exp(-i \beta \, \hat{\sigma}_z \hat{S}_y) = \hat{1} \cos\beta - i \, \hat{\sigma}_z \hat{S}_y \sin\beta \, . \tag{4.6}$$

Let the initial apparatus wave function be

$$\begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} ; \tag{4.7}$$

here we use matrix notation in the basis where the apparatus z-spin $\hat{S}_z$ is diagonal. Just after the interaction, the joint state of the particle and the apparatus becomes

$$\sum_j \begin{pmatrix} \cos(\theta+\beta_j) \\ \sin(\theta+\beta_j) \end{pmatrix} |\phi_j\rangle \langle \phi_j | \psi_1 \rangle \; , \tag{4.8}$$

where

$$\beta_j \equiv (-1)^{j+1}\beta\,, \quad j=1,2\,. \tag{4.9}$$

If the apparatus is observed to have spin up in the z-direction, then the particle's state changes according to

$$|\psi_1\rangle \rightarrow \Big(|\phi_1\rangle \cos(\theta+\beta)\langle\phi_1| + |\phi_2\rangle \cos(\theta-\beta)\langle\phi_2| \Big)\, |\psi_1\rangle\,. \tag{4.10a}$$

Similarly, if the apparatus is observed to have spin down in the z-direction, then the particle's state changes according to

$$|\psi_1\rangle \rightarrow \Big(|\phi_1\rangle \sin(\theta+\beta)\langle\phi_1| + |\phi_2\rangle \sin(\theta-\beta)\langle\phi_2| \Big)\, |\psi_1\rangle\,. \tag{4.10b}$$

These changes represent the same pure Operation as Eqs. (4.3), with the identifications $A=\cos(\theta+\beta)$ and $B=\cos(\theta-\beta)$.

Just as in d'Espagnat's model, we can obtain an ''ideal'' measurement by choosing the initial apparatus state and the coupling strength appropriately. Choose the angle $\theta=-\pi/4$ so that the apparatus spin is initially oriented along the negative x-axis, and choose the coupling angle $\beta=\pi/4$ so that the apparatus spin rotates by an angle $\pi/2$. Then, after the interaction, the apparatus z-spin is up (down) if the particle's z-spin was up (down), and the Operations represented by Eqs. (4.10) project onto the states $|\phi_1\rangle$ and $|\phi_2\rangle$.

ACKNOWLEDGMENTS

We thank Asher Peres for bringing Ref. 5 to our attention and for encouraging us to write this note. This work was supported in part by the National Science Foundation [Grant No. PHY-8620006].

REFERENCES

[1] K. Kraus, *States, Effects, and Operations: Fundamental Notions of Quantum Theory* (Springer, Berlin, 1983).

[2] G. Ludwig, *Foundations of Quantum Mechanics I and II* (Springer, Berlin, 1983).

[3] E. B. Davies, *Quantum Theory of Open Systems* (Academic, London, 1976).

[4] A. Barchielli, L. Lanz, and G. M. Prosperi, Nuovo Cimento B **72**, 79 (1982).

[5] B. d'Espagnat, Found. Phys. **16**, 351 (1986).

[6] C. M. Caves and G. J. Milburn, Caltech preprint GRP-134, May 1987, submitted to Phys. Rev. A.

CHAPTER 5

Wringing Out a Better Bell

by Samuel L. Braunstein and Carlton M. Caves

ABSTRACT

i) Usually an inequality that is chained becomes a weaker inequality. We show that chaining the Bell[1] inequality can actually lead to stronger violations by quantum mechanics. ii) By using the notion of the information gained in a measurement, we are able to derive a new Bell like inequality. This information theoretic Bell inequality automatically applies to more general situations than the standard Bell inequality (for instance, higher-spin versions of the EPR experiment), and further, it is violated when noise-free measurements yield less information (on average) than logic based on local realism would predict. iii) We point out that the Bell inequality fits naturally into a hierarchy of inequalities predictable from local realism.

I. INTRODUCTION

Local realism implies constraints on the statistics of two widely separated physical systems. These constraints, collectively known as Bell[1–3] inequalities, can be violated by quantum mechanics. The Clauser-Horne-Shimony-Holt[4] (CHSH) Bell inequality applies to a pair of two-state systems and constrains the value of a certain linear combination of four correlation functions between the two systems. Quantum mechanics violates the CHSH inequality; the violation has been confirmed experimentally.[5–7]

This chapter introduces a new kind of Bell inequality in terms of information theoretic quantities. The general formulation of this inequality allows it to be applied to many situations immediately and further leads to the realization that a hierarchy of Bell inequalities exists for multiparticle systems, for which the two-particle Bell inequalities seem to play a special role.

In Section II, we introduce the auxiliary concept of chaining the standard (CHSH) correlation Bell inequality, and we show how stronger violations can be obtained in this way. We motivate our approach of writing down information theoretic quantities in Section III; for this purpose information theory is briefly reviewed there. In Section IV, we give the derivation of the information Bell inequality and apply it to higher-spin versions of the EPR[8] experiment. Chaining of these new inequalities shows how there are violations for arbitrarily high spins even at large angles. Finally, in Section V, we discuss our thoughts on the Bell inequality.

We start with a brief review of the current experimental status of tests of quantum theory versus local realism. A set of experiments performed in the midseventies have been reviewed by Clauser and Shimony[3]. These experiments involved various kinds of sources for the pair of correlated particles: atomic cascade-producing pairs of photons, annihilation of positronium, and proton-proton scattering. However, all of these experiments were substantially improved upon in new atomic cascade experiments by Aspect and various coworkers[5–7] in the early eighties. The major improvements in their experiments are worth mentioning. They started[5] with a high-efficiency source of low-energy pairs of photons emitted in an atomic-cascade of calcium. This gave them good statistics in short runs and allowed them to have large source-polarizer separations (up to 6.5 m). The stability of their source reduced problems from drifts in the source intensity between runs with the polarizers set at different orientations. Next[6] they included two-channel polarizers, which allowed them to measure directly the polarization correlations without resorting to the assumption that photons not detected behind the polarizer were in the opposite polarization from the analyzer's setting. Subsequently,[7] they included acoustic switches in front of the polarization analyzers, which could choose the orientation of the polarization analyzer

while the photons were in flight from their source, thus ruling out any information transfer between the photons. The results of these final experiments agree with quantum mechanics and violate the CHSH inequality by five standard deviations, showing empirically that a local realistic description of two-photon decays is not possible.

It is worth mentioning two known caveats to these experiments. Since the photodetectors have a low efficiency, only a small fraction of all photon pairs are actually detected. It follows that there is an extra assumption needed: that the pairs observed actually make an unbiased sample of all pairs. It has been suggested[3] that some unknown systematic error could conceivably "enhance" the correlations in such a way as to mock quantum mechanics. The second objection[9] questions whether locality has been strictly enforced with even the largest source-polarizer separations used so far. It turns out that in the experiments performed, at least one of the photons comes from an atomic level with a lifetime exceeding 40nsec – which is longer than it takes light to travel the maximum separations used – so subluminal information exchanges are not entirely ruled out.

II. CHAINING THE CHSH BELL INEQUALITY

In this section we chain the CHSH Bell inequality [En. (II.3.3) of Chapter I] to derive a "weaker" condition for local realism. Interestingly enough, this new inequality has violations that occur over a larger set of angles than in the original form of the inequality, and further, can lead to stronger violations.

In the formulation of the EPR[10] paradox for a pair of photons formed from a decay, moving back to back in a spin-singlet state (as in the Aspect et al. experiments), measurements can be made of the polarizations of the photons along the axes $\vec{a}$ and $\vec{b}$ (for the left and right moving photons, respectively). For notational

convenience, we write these polarizations as σ_a and σ_b respectively. For a left-moving photon $\sigma_a = +1$, if the photon is polarized along $\vec{a}$, and $\sigma_a = -1$, if it is polarized perpendicular to $\vec{a}$; σ_b is defined similarly for the right moving photons.

As is shown in Chapter I, objectivity and locality – a combination called local realism – enforce the CHSH Bell inequality

$$-2 \leq C(\vec{a}', \vec{b}) + C(\vec{b}, \vec{a}) + C(\vec{a}, \vec{b}') - C(\vec{a}', \vec{b}') \leq 2 . \tag{II.1}$$

Here, $\vec{a}'$ and $\vec{b}'$ are new axes for the left- and right-moving photons, respectively, with corresponding polarizations $\sigma_{a'}$ and $\sigma_{b'}$. The correlation function $C(\vec{a}, \vec{b})$ is defined as the expectation value of the product of the polarizations σ_a and σ_b:

$$C(\vec{a}, \vec{b}) \equiv <\sigma_a \sigma_b> . \tag{II.2}$$

We will now proceed to derive the chained Bell inequality. To simplify our argument, we start by writing out the right and left hand sides of inequality (II.1) separately,

$$B(\vec{a}', \vec{b}) + B(\vec{b}, \vec{a}) + B(\vec{a}, \vec{b}') \leq B(\vec{a}', \vec{b}') , \tag{II.3}$$

$$D(\vec{a}', \vec{b}) + D(\vec{b}, \vec{a}) + D(\vec{a}, \vec{b}') \geq D(\vec{a}', \vec{b}') , \tag{II.4}$$

where

$$B(\vec{a}, \vec{b}) \equiv C(\vec{a}, \vec{b}) - 1 ,$$

$$D(\vec{a}, \vec{b}) \equiv C(\vec{a}, \vec{b}) + 1 . \tag{II.5}$$

Each of these expressions [Eqns. (II.3) and (II.4)] has a correlation on the right and a sum of similar correlations at ''interpolated'' detector settings on the left. Thus, each

of these expressions can be iterated by splitting the correlations on the left into further interpolations. This gives in general

$$B(\vec{a}_1,\vec{b}_2)+B(\vec{b}_2,\vec{a}_3)+\cdots+B(\vec{a}_{N-1},\vec{b}_N)\leq B(\vec{a}_1,\vec{b}_N)\,, \tag{II.6}$$

$$D(\vec{a}_1,\vec{b}_2)+D(\vec{b}_2,\vec{a}_3)+\cdots+D(\vec{a}_{N-1},\vec{b}_N)\geq D(\vec{a}_1,\vec{b}_N)\,, \tag{II.7}$$

where N is even [the case $N=4$ is just the standard CHSH inequality of Eqn (II.1)]. In our notation, the vectors $\vec{a}_1$, $\vec{b}_2$, $\vec{a}_3$, $\vec{b}_4$, $\cdots$, $\vec{a}_{N-1}$, $\vec{b}_N$ represent orientations for the polarization analyzers of the left- and right-moving photons in runs where these settings are taken pairwise, as in Eqns. (II.6) and (II.7). Further, the polarizations at these settings are labeled σ_1, σ_2, σ_3, σ_4, $\cdots$, σ_{N-1}, σ_N, respectively.

Expressions (II.6) and (II.7) are the chained Bell inequalities. It may be worth noting that they can be easily derived directly from the assumption of local realism without resorting to chaining. To do this, we shall follow an argument (given in Chapter I, Section II.3) similar to one used by Peres[11] for the CHSH inequality [Eqn. (II.1)]. Under the assumption of realism (objectivity), every one of these polarizations is an objective property for each decay — independent of their observation. Given this, it is sensible to write expressions for these properties of a single decay, for instance,

$$(\sigma_1\sigma_2-1)+(\sigma_2\sigma_3-1)+\cdots+(\sigma_{N-1}\sigma_N-1)\leq(\sigma_1\sigma_N-1)\,, \tag{II.8}$$

(N is even). This inequality is indeed true for each and every decay, as we see in the following way: Each polarization σ_i will be ±1, so that each bracketed term is at most zero. We have two cases: (i) σ_1 and σ_N have the same sign, so that expression (II.8) is trivially true; (ii) σ_1 and σ_N have opposite signs, in which case the right-hand side is -2, but since N is even, at least one of the pairs (σ_1, σ_2), (σ_2, σ_3), $\cdots$, (σ_{N-1}, σ_N)

must also have opposite signs, so the left-hand-side is no bigger than -2. We make use of realism once more to convert expression (II.8) into one involving measurable correlation functions. Since for each decay all the polarizations are objective, over many decays the statistics of these objective properties must form a grand probability distribution,

$$p(\sigma_1, \sigma_2, \cdots, \sigma_N)\,. \tag{II.9}$$

We don't claim to be able to know the values $\sigma_1, \cdots, \sigma_N$ for each decay, being able only to determine a pair of them in any single experimental run. However, we may use this distribution to average (II.8) over many decays. This averaging gives exactly inequality (II.6). We can derive inequality (II.7) in the same way if, instead, we start with the expression

$$(\sigma_1\sigma_2+1)+(\sigma_2\sigma_3+1)+\cdots+(\sigma_{N-1}\sigma_N+1)\geq(\sigma_1\sigma_N+1)\,, \tag{II.10}$$

which also holds good for every single decay if realism is true. For convenience, we can combine the inequalities (II.6) and (II.7) into one expression,

$$|C(\vec{a}_1,\vec{b}_2)+C(\vec{b}_2,\vec{a}_3)+\cdots+C(\vec{a}_{N-1},\vec{b}_N)-C(\vec{a}_1,\vec{b}_N)|\leq N-2\,. \tag{II.11}$$

To be able to compare the correlation functions that appear in Eqn. (II.11) in a compelling way to measured correlations, we need to assume that the measurement of a polarization at one end of the apparatus doesn't disturb the measurement of the distantly separated photon. The assumption of locality allows this. Thus, the objectivity (realism) embodied in Eqn. (II.11) becomes a rigorous expression for measured correlations under the assumption of "local realism."

This chained Bell inequality [Eqn. (II.11)] has been derived in the literature before,[12] but very little comment has been made about it. Only one theorem exists,[2] which states that any Bell inequality that is of the form of a linear combination of the correlation functions can be derived from linear combinations of the CHSH inequality. Our reasons for presenting the chained Bell inequality again are to point out some interesting experimental consequences of it, and to illustrate how the assumption of realism must be used at two points: to give an expression that is valid for individual decays, and to be able to average over this expression to obtain an inequality in terms of measurable quantities. The existence of the grand probability as being a consequence of realism is usually underplayed, but we will find it crucial in deriving our information Bell inequalities. We are not the only researchers to point out the significance of the existence of these grand probabilities.[13]

We still have to show that these chained inequalities are violated by quantum mechanics. The ideal quantum mechanical prediction for the photonic spin-singlet state is

$$C(\vec{a},\vec{b}) \equiv C(\theta) = -\cos 2\theta \,, \tag{II.12}$$

where θ is the angle between $\vec{a}$ and $\vec{b}$. The best geometry to choose is where successive vectors in the list $\vec{a}_1, \vec{b}_2, \vec{a}_3, \vec{b}_4, \cdots, \vec{a}_{N-1}, \vec{b}_N$ are separated by angle $\theta/(N-1)$ (see Fig. 1 for this geometry). With this geometry, Eqn. (II.11) shows that we get a violation when the quantity – the signal for violation –

$$S_N \equiv |(N-1)C(\theta/(N-1)) - C(\theta)| - (N-2) \tag{II.13}$$

is positive. For a small angle ϕ, Eqn. (II.12) predicts

$$C(\phi) \sim -1 - 2\phi^2 \,, \tag{II.14}$$

so when the pair of detectors is closely aligned, the outputs from the detectors are nearly exactly anticorrelated. Taking the limit of large N, at fixed θ, one finds that the signal for violation is

$$S_N \to 1 + C(\theta) = 1 - \cos(2\theta) , \tag{II.15}$$

which shows a violation at every angle except multiples of π, since perfect anticorrelation occurs only when the detectors are perfectly aligned. We see that the violation occurs because quantum mechanics is more tightly correlated than local realism allows. The more we take smaller ''slicings'' of settings between fixed outer vectors $\vec{a}_1$ and $\vec{b}_N$, the less the effect of the combined quantum correlations summed over the ''slices.'' This is analogous to the quantum Zeno paradox,[14] where it is found that repeated measurements of a system at closer and closer times ''stops'' the quantum evolution of the state.

Looking at the expression (II.15) for S_N, one might think that the violations are somewhat larger (for $N = \infty$) than for the CHSH inequality (*i.e.*, for $N = 4$). However, to compare the magnitude of violations meaningfully, we need to ask how many standard deviations a violation corresponds to; that is, we must say something about the size of the ''experimental'' error bars that will appear in the signal.

To estimate this signal to noise, we shall assume that the only noise is the statistical noise that comes from trying to estimate the correlation functions from a finite number of counts. To be able to keep all runs within a time over which the two-photon source is stable, we want to keep the total running time as short as possible. Suppose we restrict ourselves to a run time that yields on average $\mathcal{N}$ counts. For the chained Bell inequality given by the expression (II.11), we need to measure N correlation functions; we must split up the total running time suitably to measure them all.

Now the correlation function for a given pair of polarizer settings can be written as

$$C(\theta)=P_{++}(\theta)+P_{--}(\theta)-P_{+-}(\theta)-P_{-+}(\theta)\,, \tag{II.16}$$

where $P_{+-}(\theta)$ is the probability of obtaining a +1 polarization for the left-moving photon and a -1 for the right-moving one, etc. Experimentally, each probability P_i ($i = ++$, --, etc) is estimated to be the frequency $g_i = R_i/M$, where R_i is the number of counts for the pair of results specified by i, and $M = R_1+R_2+R_3+R_4$ is the total number of counts for this pair of settings. Given these estimated probabilities, one estimates the correlation function to be

$$C_{\exp}(\theta)=g_{++}(\theta)+g_{--}(\theta)-g_{+-}(\theta)-g_{-+}(\theta)$$

$$\equiv\sum_i \lambda_i g_i\,, \qquad \lambda_i=\pm 1\,. \tag{II.17}$$

For a fixed total number of counts M, the numbers of counts R_i for two channels are distributed according to the multinomial distribution

$$p(R_1,R_2,R_3,R_4\,|\,M)=\frac{M!}{R_1!R_2!R_3!R_4!}\,P_1^{R_1}P_2^{R_2}P_3^{R_3}P_4^{R_4}\,. \tag{II.18}$$

To determine the noise associated with $C_{\exp}(\theta)$, we need the covariance matrix of the frequencies of occurrence of each channel (which are the estimates for the P_i). For the number of counts in each channel, we find

$$(\overline{R_i})_{M\ \text{fixed}}=MP_i \tag{II.19}$$

$$(\overline{R_iR_j})_{M\ \text{fixed}}=M(M-1)P_iP_j+MP_i\delta_{ij}\,, \tag{II.20}$$

where the barred quantities denote a statistical average. In terms of the measured

frequencies g_i, averaged over the total number of counts M, we find the covariance matrix to be

$$c_{ij} \equiv \overline{g_i g_j} = \overline{\left[\frac{1}{M}\right]} (P_i \delta_{ij} - P_i P_j) . \tag{II.21}$$

The average of the experimentally estimated correlation function is the quantum mechanical prediction; *i.e.*,

$$\overline{C_{\exp}(\theta)} = C(\theta) , \tag{II.22}$$

and the variance of $C_{\exp}(\theta)$ is

$$\begin{aligned} \overline{[\Delta C_{\exp}(\theta)]^2} &= \sum_{i,j} \lambda_i \lambda_j \overline{g_i g_j} \\ &= \overline{\left[\frac{1}{M}\right]} [1 - C^2(\theta)] . \end{aligned} \tag{II.23}$$

We notice that this is strongly dependent on angle; near $\theta = 0$, we have an almost perfect anticorrelation of which we are very certain.

The signal for a violation of expression (II.11) will be when

$$| C(\vec{a}_1, \vec{b}_2) + C(\vec{b}_2, \vec{a}_3) + \cdots + C(\vec{a}_{N-1}, \vec{b}_N) - C(\vec{a}_1, \vec{b}_N) | - (N-2) \tag{II.24}$$

is greater than zero. The experimental estimate of this signal is

$$\begin{aligned} S_{N,\exp} \equiv | & C_{\exp}(\vec{a}_1, \vec{b}_2) + C_{\exp}(\vec{b}_2, \vec{a}_3) + \cdots \\ & + C_{\exp}(\vec{a}_{N-1}, \vec{b}_N) - C_{\exp}(\vec{a}_1, \vec{b}_N) | - (N-2) , \end{aligned} \tag{II.25}$$

with mean

$$\overline{S_{N,\exp}} \equiv S_N = |(N-1)C(\theta/(N-1)) - C(\theta)| - (N-2) . \qquad \text{(II.26)}$$

(The absolute value causes a problem in taking the mean as the absolute value approaches zero; this is a region, however, where no violation occurs and so is of no interest, so we ignore this problem.)

We have available a total time yielding on average $\mathcal{N}$ counts, and we can apportion this time among the measurements of the various correlation functions as we please. Suppose we use a fraction f of the time to measure $C(\theta)$ [we take $\overline{M} = f\mathcal{N}$ in Eqn. (II.23) and approximate $\overline{1/M}$ by $1/\overline{M}$], leaving the remaining time to measure the $N-1$ functions $C(\theta/(N-1))$ at the various settings [for these we take $\overline{M} = (1-f)\mathcal{N}/(N-1)$]. The variance of the estimated signal Eqn. (II.25) will tell us the size of the "noise":

$$N_N^2 \equiv \overline{(\Delta S_{N,\exp})^2}$$

$$= (N-1)\overline{\Delta C_{\exp}^2(\theta/(N-1))} + \overline{\Delta C_{\exp}^2(\theta)}$$

$$= \frac{1}{\mathcal{N}}\left\{ \frac{(N-1)^2}{1-f}\left[1 - C^2(\theta/(N-1))\right] + \frac{1}{f}\left[1 - C^2(\theta)\right]\right\} . \qquad \text{(II.27)}$$

Minimizing this noise with respect to the division of time will maximize the signal to noise; the minimum occurs with f specified by

$$f = \frac{\left[1 - C^2(\theta)\right]^{1/2}}{(N-1)\left[1 - C^2(\theta/(N-1))\right]^{1/2} + \left[1 - C^2(\theta)\right]^{1/2}} . \qquad \text{(II.28)}$$

In this case, the noise becomes

$$N_N = \frac{1}{\sqrt{\mathcal{N}}}\left\{ (N-1)\left[1 - C^2(\theta/(N-1))\right]^{1/2} + \left[1 - C^2(\theta)\right]^{1/2}\right\} . \qquad \text{(II.29)}$$

We shall now include one systematic effect: the loss of correlation through ''data flipping error,'' whereby the +1 and −1 channels get crossed. This loss can be modeled by

$$C(\theta) = -\eta \cos(2\theta) \tag{II.30}$$

$$\eta = (1-2q)^2 , \tag{II.31}$$

where q is the probability to flip from $\pm$ to $\mp$ for each photon, and where $-\cos(2\theta)$ is just the ideal quantum-mechanical prediction for the correlation function. The quantity $1-\eta$ represents the data-flipping error. This model of the data error is exactly what Aspect et al.[6] use. They included two contributions to η: imperfections in the polarization analyzers, and detectors with a finite solid angle, which might accept coincidences from decays with the wrong helicity. In Ref. 6, Aspect et al. achieved $\eta \sim 0.955$.

To see the effect of the data error on the procedure of chaining, let us examine what happens to Eqns. (II.26) and (II.29) for large N. When N is large, the signal becomes

$$S_N \sim [1-\eta\cos(2\theta)] - N(1-\eta) , \qquad \theta \ll 2\pi N . \tag{II.32}$$

To get violations for large N, we rely on the nearly perfect anticorrelation for small angles between the polarization analyzers [see Eqn. (II.14)]. Thus, if $\eta \neq 1$, the signal is degraded increasingly as N increases. When $\eta = 1$, however, we recover Eqn. (II.14).

For large N, the noise can be split into two cases:

i) $\eta = 1$, for which large N yields:

$$N_N \sim \frac{1}{\sqrt{\mathcal{N}}}\left[\,|\sin(2\theta)|+2|\theta|\right], \quad \theta \ll 2\pi N\ . \tag{II.33}$$

The noise, in this case, has cusps at $\theta = 0°$ and 90°. Further, the signal to noise becomes:

$$\frac{S_N}{N_N} = \sqrt{\mathcal{N}}\,\frac{2\sin^2\theta}{|\sin(2\theta)|+2|\theta|}, \quad \theta \ll 2\pi N\ , \tag{II.34}$$

which has a maximum value of $2/\pi \sim 0.6366$ at $\theta = 90°$.

ii) $\eta \neq 1$. For large N, the noise now takes the form:

$$N_N = \frac{1}{\sqrt{\mathcal{N}}}\left[N(1-\eta^2)^{1/2}+(1-\eta^2\cos^2 2\theta)^{1/2}\right], \quad \theta \ll 2\pi N\frac{\sqrt{1-\eta^2}}{\eta}\ . \tag{II.35}$$

This shows us that the noise increases as we increase the data error $1-\eta$.

From Eqns. (II.26) and (II.27), we can define a signal-to-noise for a single count $(S/N)_{\mathcal{N}=1}$, in terms of which the signal-to-noise for an experiment with a total of $\mathcal{N}$ counts becomes

$$\left[\frac{S_N}{N_N}\right]_{\mathcal{N}} \equiv \sqrt{\mathcal{N}}\left[\frac{S}{N}\right]_{\mathcal{N}=1}\ . \tag{II.36}$$

The single count signal-to-noise ratio $(S/N)_{\mathcal{N}=1}$ gives the size of the violation in units of statistical noise. In Fig. 2(a)-(c) we have plotted $(S/N)_{\mathcal{N}=1}$ for the cases $N = 4$, 6, 8 and 20, when (a) $\eta = 1.0$, (b) $\eta = 0.98$, and (c) $\eta = 0.955$. For $\eta = 1$ [Fig. 2(a)], the size of the peak violation increases for increasing N, there is a noticeable cusp at $\theta = 90°$ that comes from the cusp in the noise for $\eta = 1$ [see Eqn. (II.33)]. In practice, we cannot go on increasing N forever. Even before we get to the point where we can no longer measure the angles accurately, other sources of noise will become a problem.

For instance, once we introduce some data error, it is no longer advantageous to keep increasing N without bound. This is because the signal grows like $-N(1-\eta)$ [see Eqn. (II.32)], because of the imperfect anticorrelation at $\theta=0°$, which kills any violation. At $\eta=0.955$, which is an achievable[6] data error, we see from Fig. 2(c) that a better signal-to-noise ratio could be achieved if either the $N=6$ or $N=8$ versions of the chained Bell inequality were tested instead of the CHSH inequality (for which $N=4$). In fact, there is a 20% improvement in the signal-to-noise ratio for using the $N=6$ scheme over a test required to show violations of the CHSH Bell inequality.

III. INFORMATION THEORY AND REALISM

Our motivation for introducing information theory follows a very simple line of reasoning. If objectivity is true, then each particle must carry with it the details of the results of any particular measurement that might be made on it. All these details would constitute a lot of information – much more than the quantum system seems to carry. We will show in Section IV that in even the most thrifty realistic theory, each particle carries (on average) more information than in quantum theory, for EPR type experiments, and that this notion can be made precise through an "information Bell inequality."

We shall devote the rest of this Section to a review of information theory.[15] Consider two measurable quantities (observables) A and B, and label the (discrete) possible values of A and B by a and b. (Throughout we shall denote an observable by a capital letter and label its possible values by the corresponding lower-case letter.) Based on one's knowledge about A and B, one assigns a joint probability $p(a,b)$ for values a and b. Bayes's theorem,

$$p(a,b)=p(a\mid b)p(b)=p(b\mid a)p(a)\,, \tag{III.1}$$

relates the joint probability to conditional probabilities.

The information obtained when one discovers values a and b for A and B is

$$I(a,b) \equiv -\log p(a,b) . \tag{III.2}$$

The base of the logarithm determines the units of the information (base 2 for bits, base e for nats). In the same way,

$$I(b) \equiv -\log p(b) \tag{III.3}$$

is the information obtained when one discovers value b for B, and

$$I(a \mid b) \equiv -\log p(a \mid b) \tag{III.4}$$

is the further information obtained when one discovers value a for A, provided one already knows the value b for B. Bayes's theorem, rewritten in terms of information, becomes

$$I(a,b) = I(a \mid b) + I(b) = I(b \mid a) + I(a) . \tag{III.5}$$

A crucial role is played by the mean information obtained when finds values for A and B:

$$H(A,B) = \sum_{a,b} p(a,b) I(a,b) . \tag{III.6}$$

This mean information is the entropy of the probability $p(a,b)$. It can also be thought of as the total information carried by the quantities A and B, defined relative to one's knowledge about A and B incorporated in the probability $p(a,b)$. In the same way,

$$H(B)=\sum_b p(b)I(b) \tag{III.7}$$

is the information carried by B, and

$$H(A \mid b)=\sum_a p(a \mid b)I(a \mid b) \tag{III.8}$$

is the information carried by A, given the value b of B. It is useful to average $H(A \mid b)$ over B to obtain a conditional information carried by A,

$$H(A \mid B)=\sum_b p(b)H(A \mid b)=\sum_{a,b} p(a,b)I(a \mid b)\,. \tag{III.9}$$

An immediate consequence of Bayes's theorem is the relation

$$H(A,B)=H(A \mid B)+H(B)=H(B \mid A)+H(A)\,. \tag{III.10}$$

We require one more ingredient, the mutual information

$$I(a;b)\equiv I(a)-I(a \mid b)=I(b)-I(b \mid a)=I(b;a)\,. \tag{III.11}$$

This mutual information can be either positive or negative, but its mean,

$$H(A;B)=\sum_{a,b} p(a,b)I(a;b)=\sum_b p(b)\left[\sum_a p(a \mid b)\log\frac{p(a \mid b)}{p(a)}\right]\geq 0\,, \tag{III.12}$$

is nonnegative. The mean mutual information,

$$H(A;B)=H(A)-H(A \mid B)=H(B)-H(B \mid A)=H(B;A)\,, \tag{III.13}$$

is the information carried in common (mutually) by A and B.

The results we need from information theory are the two inequalities

$$H(A \mid B) \leq H(A) \leq H(A, B) . \tag{III.14}$$

The left-hand inequality, a consequence of the nonnegativity of $H(A ; B)$, means that removing a condition never decreases the information carried by a quantity. The right-hand inequality, a consequence of Eqn. (III.10), means that two quantities never carry less information than each quantity carries separately.

IV. INFORMATION BELL INEQUALITY

To derive a Bell inequality for information, we must start with a statement that embodies "classical logic" (*i.e.*, realism). Since realism demands that the photons' polarizations must be objective quantities, we may, as a consequence, write down a grand probability distribution for the polarization at every angle. For the standard Bell type experiment with two settings for each detector, this reduces to the assumption that the probability

$$p(\sigma_{a'}, \sigma_b, \sigma_a, \sigma_{b'}) \tag{IV.1}$$

exists; this probability was also necessary for the correlation Bell inequality as in Eqn. (II.1). Our reason for introducing this grand probability is to be able to quantify the amount of information learned from a measurement.

We shall use A, A', B, and B' to label the observables for the polarizations that take the values σ_a, $\sigma_{a'}$, σ_b, and $\sigma_{b'}$, respectively. An obvious generalization of the

right-hand inequality (III.14) yields

$$H(A,B) \leq H(A,B',A',B)$$

$$= H(A \mid B',A',B) + H(B' \mid A',B) + H(A' \mid B) + H(B)\,, \qquad \text{(IV.2)}$$

where we use Eqn. (III.10) to expand out the right-hand side. The right-hand side involves probabilities of noncommuting observables and hence would not be defined in quantum mechanics. We can, however, use a slight generalization of the left-hand side of inequality (III.14) to eliminate conditions:

$$H(A \mid B',A',B) \leq H(A \mid B') \qquad \text{(IV.3)}$$

$$H(B' \mid A',B) \leq H(B' \mid A')\,. \qquad \text{(IV.4)}$$

Subtracting $H(B)$ from both sides of Eqn. (IV.2), we obtain the information Bell inequality

$$H(A \mid B) \leq H(A \mid B') + H(B' \mid A') + H(A' \mid B)\,. \qquad \text{(IV.4)}$$

The four pieces of conditional information in this Bell inequality involve pair probabilities that are defined in quantum mechanics; they can be determined from the statistics of the four types of experimental runs. Note the very similar form of this inequality with that of the CHSH inequality [see Eqns. (II.1), (II.3) and (II.4)].

In Eqn. (IV.1), we made use of the assumption of reality. Where does the assumption of locality come in? [In the same way it did for the correlation Bell inequalities of Section II.] In order to relate the two-particle probabilities $p(\sigma_a, \sigma_b)$ and $p(\sigma_a \mid \sigma_b)$ to actual measured quantities, we must be able to assume that the detection of one photon's polarization doesn't affect the detection of the other's. The locality

assumption is coming in only to allow us to make an unambiguous connection between the expressions determined by classical logic and experiment.

What does quantum mechanics predict? Again, it turns out that the geometry of Fig. 1 (with $N = 4$) gives the largest violations. By spherical symmetry, we can restrict our attention to the quantity

$$\mathcal{H}(\theta) \equiv 3\, H(\theta/3) - H(\theta)\,. \tag{IV.5}$$

There will be a violation of the information Bell inequality Eqn. (IV.4), whenever $\mathcal{H}(\theta)$ is negative. When this occurs, it is because there is a deficit of information at small angles, compared with what local realism predicts. The size of this deficit is $\mathcal{H}(\theta)$ and is measured in bits.

For the case of the photonic spin-singlet EPR experiment we have been analyzing, the quantum-mechanical probabilities are

$$p(\sigma_a = +1, \sigma_b = +1) = p(\sigma_a = -1, \sigma_b = -1) = \frac{1}{2}\sin^2\theta\,,$$

$$p(\sigma_a = +1, \sigma_b = -1) = p(\sigma_a = -1, \sigma_b = +1) = \frac{1}{2}\cos^2\theta\,, \tag{IV.6}$$

which from Eqn. (III.9) gives

$$H(\theta) = -\sin^2\theta \log \sin^2\theta - \cos^2\theta \log \cos^2\theta\,. \tag{IV.7}$$

We have plotted $\mathcal{H}(\theta)$ [see Eqn. (IV.5)] in Fig. 3 for the spin one-half case discussed below. We can translate this figure into the relevant one for the photonic case by reading half the value of the angle on the abscissa of the plot. We see from this plot that a violation of Eqn. (IV.4) is predicted for the range of angles $0 < \theta < 43°$, with a maximum information deficit of 0.24 bits at $\theta = 26.6°$.

Since information is such a general concept, we can apply our information Bell inequality to many situations (the variables A, B, etc, will, instead of labeling different polarizations, label more general variables of interest). Thus, we may immediately apply our new inequality, expression (IV.4), to, for instance, higher-spin versions[16] of the EPR experiment.

We shall consider a spin-singlet state breaking into two pieces, each of spin s (in units of $\hbar$). Each piece has its spin projection measured when the two are separated by a great distance. The spin projections now take on the $2s+1$ values $-s, \cdots, s$. To see that this inequality [Eqn. (IV.4)] is actually violated by quantum mechanics, we choose the geometry as in Fig. 1 and calculate the probability of finding particle $\mathcal{A}$ in the projection $\sigma_a = m_a$ (along the z-axis) for a detector at angle θ, and particle $\mathcal{B}$ with $\sigma_b = m_b$ at zero angle:

$$p(\sigma_a = m_a, \{\theta\}; \sigma_b = m_b) = |\, {}_{\mathcal{A}}\!<s, m_a| \otimes {}_{\mathcal{B}}\!<s, m_b|\, \hat{R}_{\mathcal{A}}(\theta)\, |0,0>|^2$$

$$= \frac{|d_{m_a,-m_b}(\theta)|^2}{2s+1}\,. \tag{IV.8}$$

Here, the two-particle, spin zero state is given by[16]

$$|0,0> = \frac{1}{(2s+1)^{1/2}} \sum_{m=-s}^{s} (-1)^{s-m}\, |s, m>_{\mathcal{A}} \otimes |s, -m>_{\mathcal{B}}\,, \tag{IV.9}$$

$\hat{R}_{\mathcal{A}}(\theta)$ rotates the state of particle $\mathcal{A}$ by an angle θ around the y-axis, and $d_{m,m'}(\theta)$ is one of the matrix elements of this rotation operator between spin projections m and m'. From Eqns. (III.8) and (IV.8), we may calculate the average information

$$H(\theta) = -\sum_{m} \frac{|d_{m,-m}(\theta)|^2}{2s+1} \log |d_{m,-m}(\theta)|^2\,. \tag{IV.10}$$

We have calculated the information difference $\mathcal{H}(\theta)$ [see Eqn. (IV.5)] from the average information for the $s = 1/2, 1, 2, 5$, and 25, and plotted them on Fig. 3. We can see from Fig. 3 that the size of the violation (in terms of the information deficit) increases with increasing spin.

In the first part of this paper, we discussed how the standard Bell inequality could be chained to give stronger violations. Well, our information inequality has the same form: The information you can learn from the detectors set at some angle is bounded from above by what you could learn from a bunch of intermediate measurements. If we chain our information inequality, we get

$$H(A_1|B_2) + H(B_2|A_3) + \cdots + H(A_{N-1}|B_N) \geq H(A_1|B_N)\,, \tag{IV.11}$$

for N even, with A_1, B_2, etc. being the observables for the two-particle system. This chaining increases the size of the violation over the original version of our information inequality. Finally, we note that this chained information inequality can be derived directly from the existence of the grand probability distribution,

$$p(\sigma_1, \sigma_2, \cdots, \sigma_N)\,, \tag{IV.12}$$

without resorting to chaining, by using an obvious extension of the argument that led to Eqn. (IV.4).

V. FOR WHOM THE BELL INEQUALITY TOLLS

It seems worthwhile to try to tie down the role of the Bell inequality. Although objective properties seem to be what is being tested, there exist simpler yet compelling arguments against them. For instance, Feynman[17] has considered a two-slit experiment in which electrons form an interference pattern. He showed that *no*

probabilities

prob (that the electron arrives at x on the screen | that it passed through slit 1)

(V.1)

and

prob (that the electron arrives at x on the screen | that it passed through slit 2)

(V.2)

are consistent with getting an interference pattern; *i.e.*, there are no objective paths. Similarly, there can be no way for the electron to carry with it a property telling it where it will end up on the screen, for if this were the case, we could go to the plane of the slits and there would be a property that told the electron which slit it had gone through, which is what was initially argued to be impossible.

So what's so special about the Bell inequalities? Well, the probabilities [Eqns. (V.1) and (V.2)] assumed by Feynman are of noncommuting quantities and so cannot exist in quantum mechanics (more importantly, they can't be measured), so it seems not unreasonable that the approaches of quantum theory and objectivity should give different answers. In fact, one might have expected this always to be the case: that objectivity postulates probabilities that don't even make sense in quantum mechanics, and so they can never be directly tested. That is, we are saying that Feynman's argument is an indirect test of objectivity, since the probabilities of Eqns. (V.1) and (V.2) are never directly measured – only talked about.

The problem is that if we wish to measure probabilities like those of Eqns. (V.1) and (V.2) (say by blocking either one slit or the other), we must modify the experimental apparatus at the slits so drastically that we disturb evolution of the wave packet leading to the interference pattern. Thus, we can never be absolutely certain that the inconsistency that Feynman pointed out isn't due to our having disturbed the electron's final position on the screen through our measurement of which slit it went through.

It is surprising that the standard Bell inequalities lead to a test of objectivity that involves only pairs of commuting observables. Thus, the Bell inequalities allow a direct test of the probabilities involved, while the Feynman argument does so only indirectly. It is perhaps worth considering this contrast between the Feynman argument and Bell inequalities in more detail. Information Bell inequalities provide a natural vehicle for such a consideration.

Bell inequalities come from assuming the existence of some grand probability for a set of quantities that do not commute in quantum mechanics. One derives from the grand probability – using classical probability logic – an inequality (for correlation functions, information, etc.) and shows that for some quantum state, the inequality is violated. Thus, strictly speaking, what all Bell inequalities test is whether some aspect of the quantum statistics can be derived from a grand probability. This is very close to the point of view promoted by Garg and Mermin,[13] who start with some marginal probabilities predicted by quantum mechanics and show that they cannot be derived from higher-order probabilities (they call this the Clauser-Horne problem).

The advantage of the information Bell inequalities is that they provide a straightforward, yet general, framework for addressing these questions. An information Bell inequality can be derived in a straightforward way for any system. In contrast, a

correlation Bell inequality (such as the CHSH inequality) requires a separate derivation for each new quantum system. Of course, information Bell inequalities do not ferret out all the weird quantum behavior. To do that, the most general approach is the Garg-Mermin approach. Information Bell inequalities are, nonetheless, useful, because they are mathematically quite straightforward compared to a Garg-Mermin test.

One would like to think that the Bell inequalities test something more cosmic than just the existence of some grand probability. Thus, one tries to attach an *interpretation* to the Bell inequalities. The first part of the interpretation is objectivity (or realism). If the quantities dealt with are objective, then they have definite values, independent of observation. We do not know these actual values, but what knowledge we do have is incorporated in a probability – the grand probability. Thus, the existence of the grand probability is interpreted as a test of objectivity.

If this were the whole story, then all Bell inequalities would be on the same footing. One must make further assumptions, however, to relate the grand probability to statistics of actual measurements. One can imagine that measurements of one or more of the quantities so disturb the system that the statistics of remaining quantities are no longer those that would be inferred from the grand probability. Indeed, this is how objectivity is maintained in a naive interpretation of quantum mechanics that goes back to Heisenberg's original discussion of the uncertainty principle. How compelling the further assumptions are determines how convincing the Bell inequality is as a test of objectivity.

Consider a single particle and two measurable quantities (observables) A and A' (which do not commute in quantum mechanics). The trivial one-particle information Bell inequality,

$$H(A) \geq H(A \mid A') \tag{V.3}$$

[Eqn. (III.14)], comes from assuming a joint probability $p(a, a')$. Examples include (i) the two-slit experiment with A' = (position at screen with slits) $= \pm 1$ and A = (discretized position at the detecting screen) and (ii) a spin-1/2 system with A' = (y-spin) and A = (z-spin). Feynman claims that the two-slit experiment contains the essence of quantum mechanics, so one might think that this Bell inequality is enough. Nonetheless, as we have argued, there is a serious objection to this: Any attempt at measuring the probabilities needed in Eqn. (V.3) by necessity involves disturbing the system before all the probabilities can be measured. In this case, the premises for inequality (V.3) do not hold.

To overcome this objection, one needs a Bell inequality where all the probabilities that appear are defined in quantum mechanics – *i.e.*, they are probabilities for commuting quantities – without having to specify a measuring procedure that might "disturb" the system. If one tries to modify Eqn. (V.3) by including an apparatus that measures A', then one finds that Eqn. (V.3) is always satisfied because the joint probability *does* exist in quantum mechanics. Indeed, as discussed above, one might have thought that in *any* case where all the probabilities are defined in quantum mechanics, there could be no violation of an inequality derived using classical probability theory. The significance of the two-system Bell inequalities is precisely that they do lead to such a violation.

The one-system Bell inequalities are not compelling because of the necessity of disturbing the system before all the information can be obtained to check them. The two-system Bell inequalities do not need to suffer from this problem, because they can be phrased solely in terms of commuting quantities. But why stop here? Surely there are three-system (and higher) Bell inequalities. We could, in principle, consider three

body decays for which the various properties of each particle would be assumed to be objective. The existence of a suitable grand probability would follow just as the one in Eqn. (II.9) did. We could even include the assumption of locality and perhaps use our information Bell inequality techniques to find an appropriate inequality. At the moment, this generalization would appear to lead to a complicated mess (especially experimentally), with no obvious gain over tests of the two-particle Bell inequalities – though it is not inconceivable that some special correlated many-body states could in the future yield a simplified test of local realism. Thus, although we recognize that there will be a whole hierarchy of Bell inequalities for many-particle systems, at the moment it appears that those for two-particles yield the simplest and most compelling tests of local realism.

ACKNOWLEDGEMENTS

One of the authors (S.L.B.) appreciated discussions with Golda Bernstein on Feynman's argument.

REFERENCES

[1] J. S. Bell, Physics **1**, 195 (1964).

[2] F. Selleri and G. Tarrozzi, Riv. Nuovo Cim. **4**, 1 (1981).

[3] J. F. Clauser and A. Shimony, Rep. Prog. Phys. **41**, 1881 (1978).

[4] J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[5] A. Aspect, P. Grangier and G. Roger, Phys. Rev. Lett. **47**, 460 (1981).

[6] A. Aspect, P. Grangier and G. Roger, Phys. Rev. Lett. **49**, 91 (1982).

[7] A. Aspect, J. Dalibord and G. Roger, Phys. Rev. Lett. **49**, 1804 (1982).

[8] For the spin-1/2 version, see D. Bohm and Y. Aharonov, Phys. Rev. **108**, 1070 (1957).

[9] S. Pascazio and J. Reignier, Phys. Lett. A **126**(3), 163 (1987).

[10] A. Einstein, N. Rosen and B. Podolsky, Phys. Rev. **47**, 777 (1935).

[11] A. Peres, Am. J. Phys. **46**, 745 (1978).

[12] P. M. Pearle, Phys. Rev. D **2**, 1418 (1970).

[13] A. Garg and N. D. Mermin, Found. Phys. **14**, 1 (1984).

[14] See, for instance, A. Perez, Am. J. Phys. **48**, 931 (1980).

[15] See, for instance, S. Goldman, *Information Theory* (Prentice-Hall, New York, 1955).

[16] N. D. Mermin, Phys. Rev. D **22**, 356 (1980).

[17] R. P. Feynman, R. B. Leighton and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesely, Reading, Massachusetts, 1977) vol. III, chapter 1.

FIGURE CAPTIONS

FIG. 1. Plot of the geometry leading to maximal violation for the chained Bell inequality, with angles $\theta/(N-1)$ between the successive vectors $(\vec{a}_1, \vec{b}_2)$, $(\vec{b}_2, \vec{a}_3)$, etc and angle θ between the outermost pair $(\vec{a}_1, \vec{b}_N)$.

FIG. 2. Plot of the "single-count signal to noise" $(S/N)_{N=1}$ versus angle, to show the violation of the chained correlation Bell inequality of Eqn. (II.11), with chainings $N = 4$, 6, 8 and 20 for the cases: (a) $\eta = 1.0$, (b) $\eta = 0.98$, and (c) $\eta = 0.955$.

FIG. 3. Plot of the information difference $\mathcal{H}(\theta)$ (in bits) versus angle, for the spins $s = 1/2$, 1, 2, 5 and 25. A violation of the information Bell inequality Eqn. (IV.4) occurs whenever $\mathcal{H}(\theta)$ is negative.

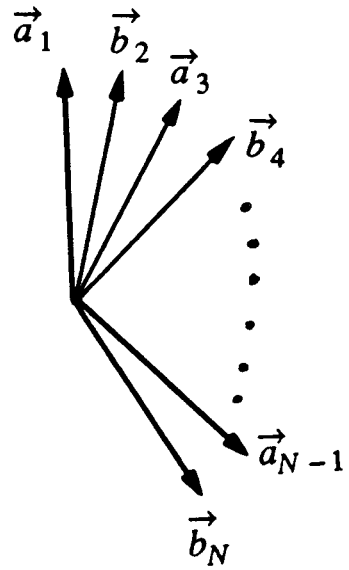

Figure 1

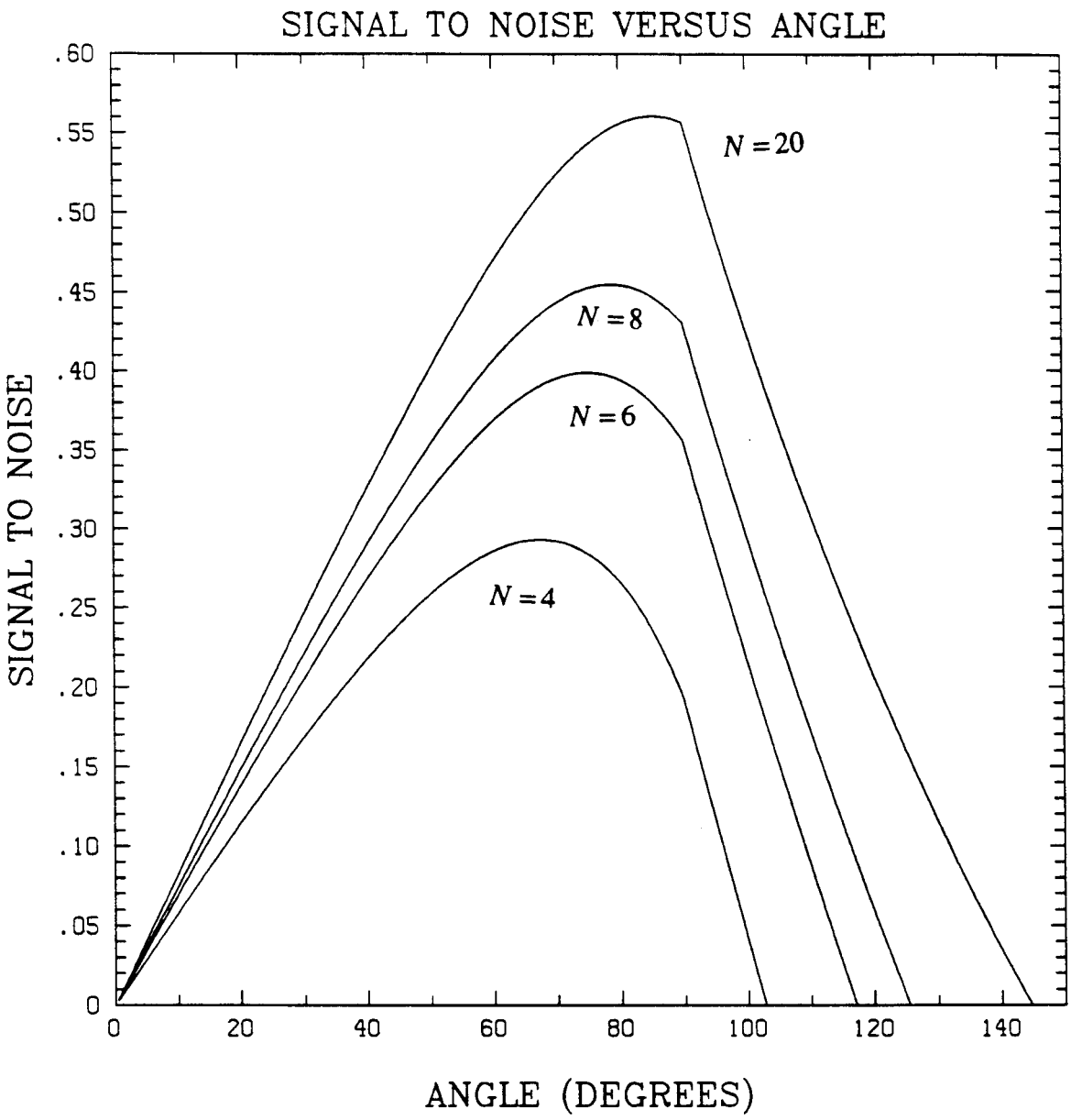
SIGNAL TO NOISE VERSUS ANGLE
SIGNAL TO NOISE
.60
.55
.50
.45
.40
.35
.30
.25
.20
.15
.10
.05
0
0
20
40
60
80
100
120
140
ANGLE (DEGREES)
N = 20
N = 8
N = 6
N = 4

Figure 2(a)

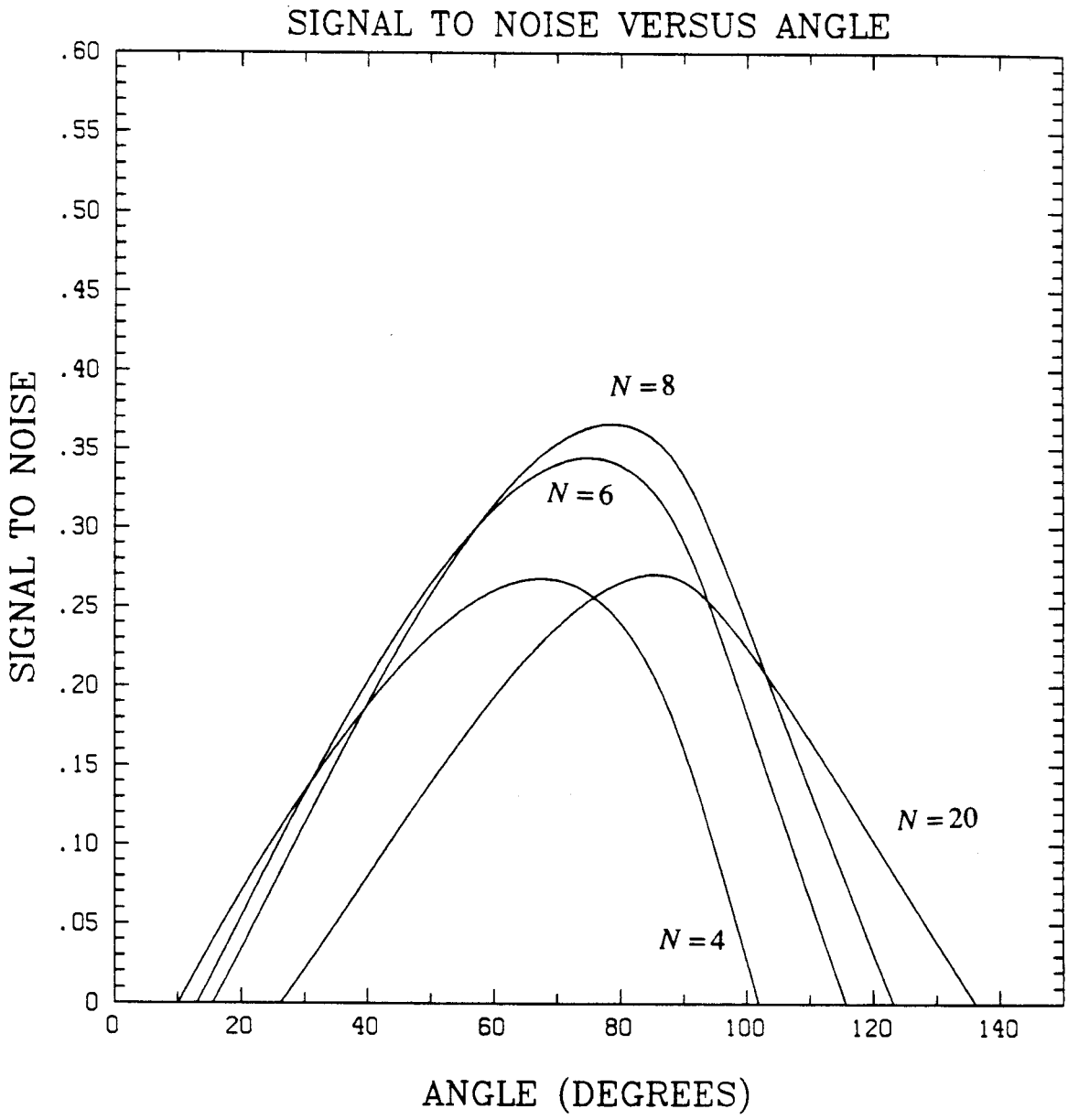
SIGNAL TO NOISE VERSUS ANGLE
SIGNAL TO NOISE
.60
.55
.50
.45
.40
.35
.30
.25
.20
.15
.10
.05
0
0
20
40
60
80
100
120
140
ANGLE (DEGREES)
N = 8
N = 6
N = 20
N = 4

Figure 2(b)

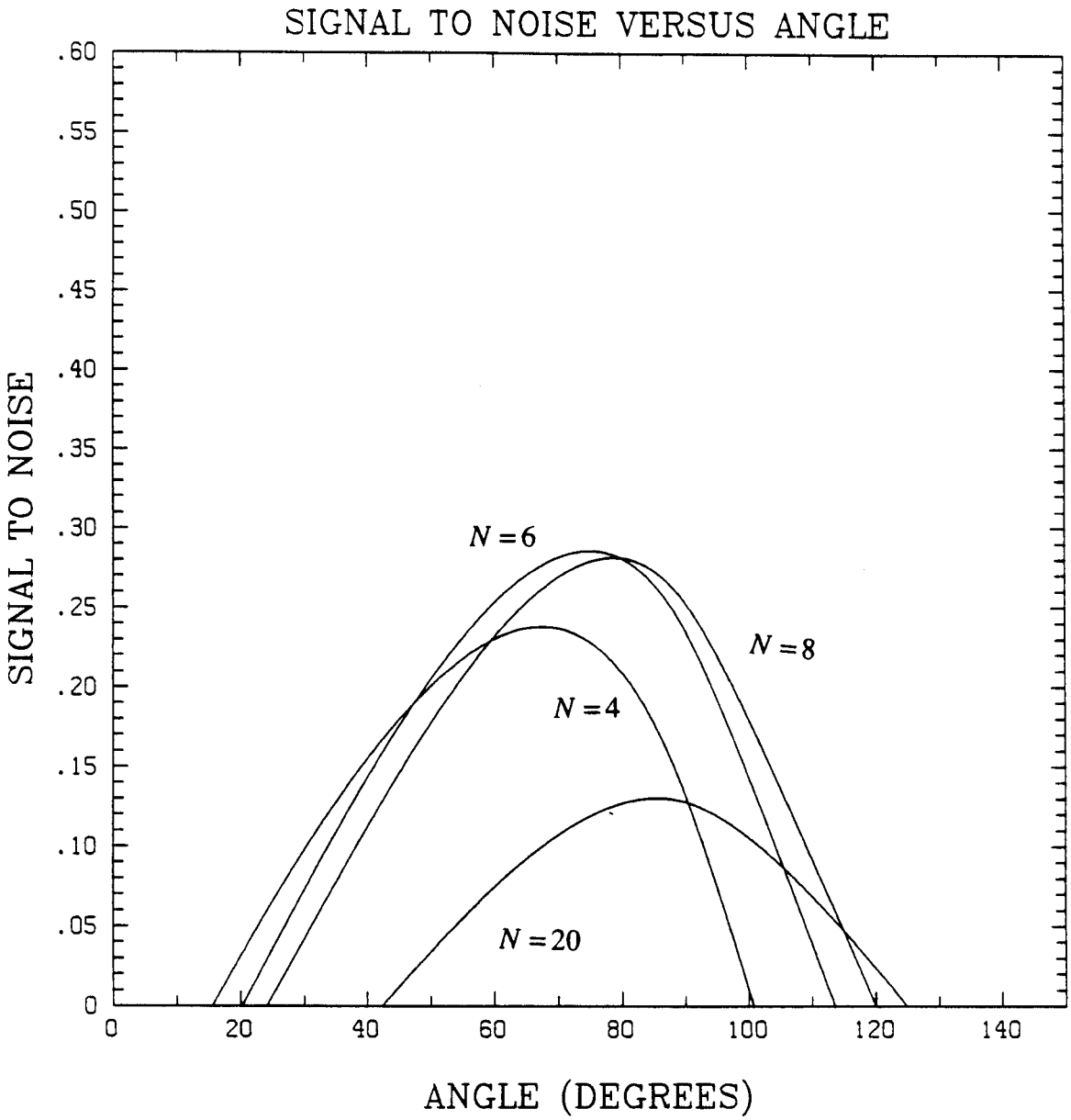
SIGNAL TO NOISE VERSUS ANGLE
SIGNAL TO NOISE
.60
.55
.50
.45
.40
.35
.30
.25
.20
.15
.10
.05
0
0
20
40
60
80
100
120
140
ANGLE (DEGREES)
N = 6
N = 8
N = 4
N = 20

Figure 2(c)

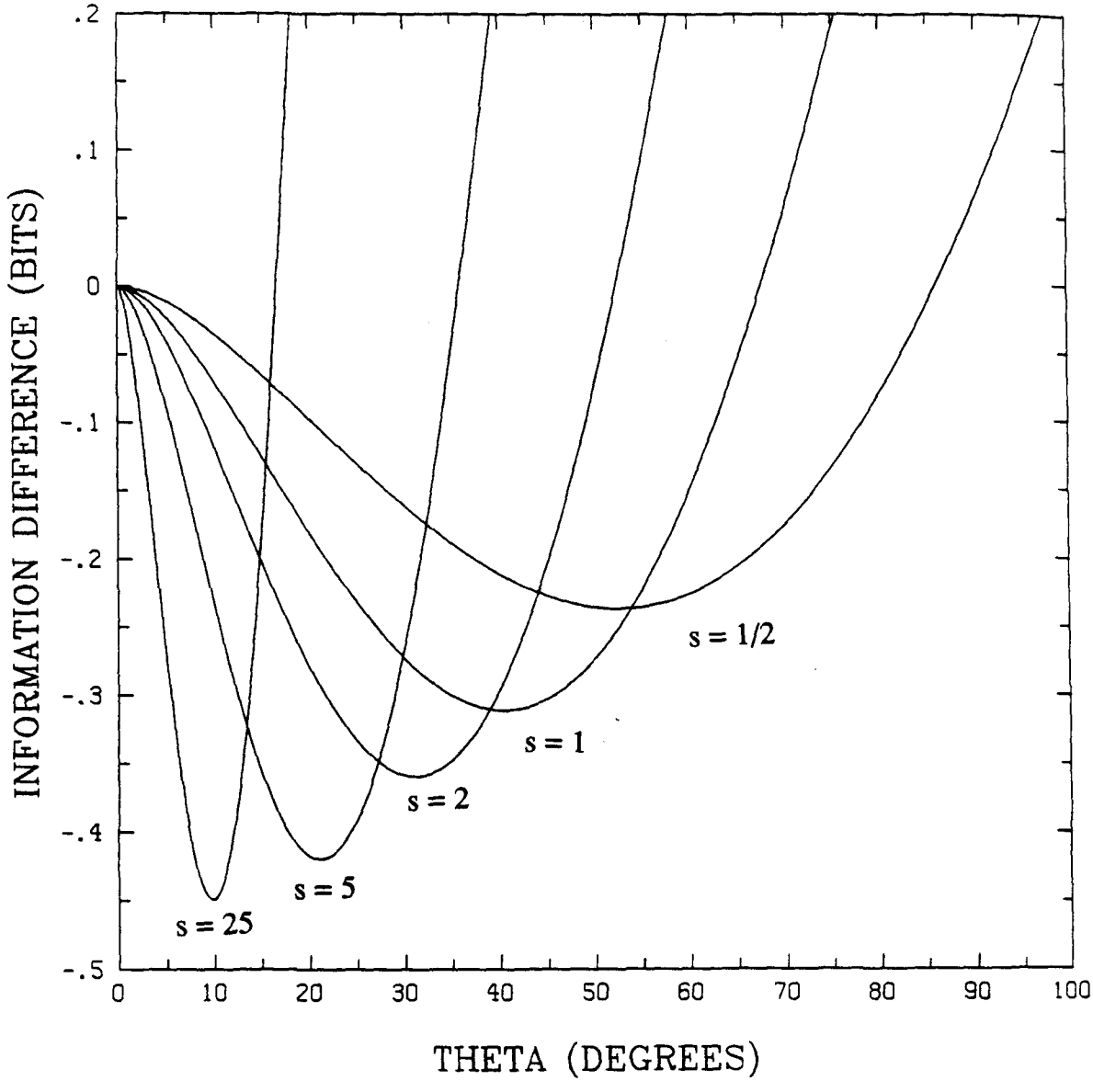
INFORMATION DIFFERENCE (BITS)
.2
.1
0
-.1
-.2
-.3
-.4
-.5
0
10
20
30
40
50
60
70
80
90
100
THETA (DEGREES)
s = 1/2
s = 1
s = 2
s = 5
s = 25

Figure 3