On the measurement of a weak classical force coupled to a quantum-mechanical oscillator. I. Issues of principle*

Carlton M. Caves, Kip S. Thorne, Ronald W. P. Drever,† Vernon D. Sandberg,‡ and Mark Zimmermann§

W. K. Kellogg Radiation Laboratory, California Institute of Technology, Pasadena, California 91125

The monitoring of a quantum-mechanical harmonic oscillator on which a classical force acts is important in a variety of high-precision experiments, such as the attempt to detect gravitational radiation. This paper reviews the standard techniques for monitoring the oscillator, and introduces a new technique which, in principle, can determine the details of the force with arbitrary accuracy, despite the quantum properties of the oscillator. The standard method for monitoring the oscillator is the "amplitude-and-phase" method (position or momentum transducer with output fed through a narrow-band amplifier). The accuracy obtainable by this method is limited by the uncertainty principle ("standard quantum limit"). To do better requires a measurement of the type which Braginsky has called "quantum nondemolition." A well known quantum nondemolition technique is "quantum counting," which can detect an arbitrarily weak classical force, but which cannot provide good accuracy in determining its precise time dependence. This paper considers extensively a new type of quantum nondemolition measurement—a "back-action-evading" measurement of the real part X_1 (or the imaginary part X_2) of the oscillator's complex amplitude. In principle X_1 can be measured "arbitrarily quickly and arbitrarily accurately," and a sequence of such measurements can lead to an arbitrarily accurate monitoring of the classical force. The authors describe explicit *Gedanken* experiments which demonstrate that X_1 can be measured arbitrarily quickly and arbitrarily accurately. In these experiments the measuring apparatus must be coupled to both the position (position transducer) and the momentum (momentum transducer) of the oscillator, and both couplings must be modulated sinusoidally. For a given measurement time the strength of the coupling determines the accuracy of the measurement; for arbitrarily strong coupling the measurement can be arbitrarily accurate. The "momentum transducer" is constructed by combining a "velocity transducer" with a "negative capacitor" or "negative spring." The modulated couplings are provided by an external, classical generator, which can be realized as a harmonic oscillator excited in an arbitrarily energetic, coherent state. One can avoid the use of two transducers by making "stroboscopic measurements" of X_1, in which one measures position (or momentum) at half-cycle intervals. Alternatively, one can make "continuous single-transducer" measurements of X_1 by modulating appropriately the output of a single transducer (position or momentum), and then filtering the output to pick out the information about X_1 and reject information about X_2. Continuous single-transducer measurements are useful in the case of weak coupling. In this case long measurement times are required to achieve good accuracy, and continuous single-transducer measurements are almost as good as perfectly coupled two-transducer measurements. Finally, the authors develop a theory of quantum nondemolition measurement for arbitrary systems. This paper (Paper I) concentrates on issues of principle; a sequel (Paper II) will consider issues of practice.

CONTENTS

*Supported in part by the National Aeronautics and Space Administration (NGR 05-002-256 and a grant from PACE) and by the National Science Foundation (Grant No. AST76-80801 A02).

†Also at Department of Natural Philosophy, Glasgow University, Glasgow, Scotland.

‡Chaim Weizmann Research Fellow.

§Robert A. Millikan Fellow.

I. INTRODUCTION

A. Message of this paper in brief

Consider a very classical incoming signal—i.e., a signal carried by a boson field with occupation number (number of quanta per quantum-mechanical state) huge compared to unity. The signal is coupled weakly to a quantum-mechanical harmonic oscillator—so weakly in fact that, if the oscillator is initially unexcited, the signal can deposit into it an average of only a few quanta per cycle; perhaps even much less than one. The objective is to measure the time dependence of the incoming signal by monitoring some aspect of the oscillator's motion. *Question*: With what accuracy can the signal be measured? *Answer*: With arbitrary accuracy, in principle. As long as one concerns oneself only with limitations imposed by nonrelativistic quantum mechanics, and as long as the signal is arbitrarily classical, then no matter how weak the coupling of the signal to the oscillator may be, it can be measured arbitrarily accurately.

However, to obtain good accuracy when the coupling is weak, one must *not* monitor the oscillator's state using currently standard electronic methods. Those methods ask the oscillator "What is your amplitude *and* phase of oscillation?"—and because amplitude and phase are noncommuting observables, the uncertainty principle forbids a precise answer. For such "amplitude-and-phase" methods the amplitude error, expressed in terms of the number of oscillator quanta N, always exceeds $(\Delta N)_{\min} = (N+\frac{1}{4})^{1/2}$; the phase error (for large N) always exceeds $(\Delta\psi)_{\min} = \frac{1}{2}N^{-1/2}$ (Serber and Townes, 1960; Braginsky, 1970; Giffard, 1976). These errors prevent accurate measurement of the incoming signal, and prevent any measurement at all in the case of weak signals.

To measure the signal more accurately, one must ask the oscillator for less information about itself—"less is more"![1] Specifically, one must ask the oscillator for the value of only one observable, and it must be an observable whose future values are precisely predictable (in the absence of forces) from the result of an initial, precise measurement. The signal is then detected by the changes it produces in the values of this observable.

A well known technique of this type is "quantum counting." This technique asks the oscillator, "How many quanta N do you have in yourself?—But don't tell me anything about your phase." In principle the query can be repeated over and over again, and the answers can be completely precise and predictable (no uncertainty!) in the absence of external forces. When $N \gg 1$, quantum counting can reveal, in principle, an incoming signal far weaker than those detectable by the "amplitude-and-phase" method. However, it cannot detect signals so weak as to change N by less than unity; and for strong signals, it cannot measure the signal strength more precisely than a factor ~3 (cf. Sec. II.D).

Recently, the authors of this article have proposed new methods of measurement (Thorne *et al.*, 1978, 1979). In these methods one says to the oscillator, "What is the real part X_1 of your complex amplitude?—But don't tell me anything about the imaginary part X_2." In principle, the query can be repeated as often as desired, the answers can come through with arbitrary accuracy, and they can lead to an arbitrarily accurate monitoring of an arbitrarily weak, classical incoming signal. We call such measurements "back-action-evading" because they permit the real part of the complex amplitude X_1 to evade the back-action forces of the measuring apparatus (at the price of increasing the back-action forces on the imaginary part X_2).[2]

---

[1] "Less is more" is an aphorism popularized in this century by architect Ludwig Mies van der Rohe. It appears earlier in Robert Browning's poem *Andrea del Sarto* (1855), l. 78. Wyler (1974) has used "less is more" and related ideas as fundamental conceptual tools for exploring the frontiers of modern physics and cosmology.

[2] Hollenhorst (1979) has pointed out that back-action-evading measurements of X_1 are analogous to sending an electronic signal through a degenerate parametric amplifier. Such an amplifier takes the input signal $\mathrm{Re}[(V_1 + iV_2)e^{-i\omega t}]$ from an *ideal* voltage source, and preferentially amplifies the real part of the complex amplitude while attenuating the imaginary part; the amplifier's output is $AV_1 \sin\omega t - (V_2/A)\cos\omega t$ (Takahasi, 1965). While this is *analogous* to a back-action-evading measurement of the real part X_1 of the complex amplitude $X_1 + iX_2$ of an oscillator, it is by no means *the same*. For example, if one simply attaches a capacitive position transducer to a mechanical oscillator, and follows it by a degenerate parametric amplifier, the amplifier will act back on the oscillator through the transducer to *drive directly* the X_1 oscillations which it seeks to measure. Such a measurement is not back-action-evading. On the other hand, by a clever nonstraightforward use of degenerate parametric amplification, one *can* perform back-action-evading measurements of changes in X_1 [Hollenhorst (private communication); see Paper II]. For comments on the related issue of "phase-sensitive detection" and its relationship to "back-action evasion," see Sec. II.F.3.

Braginsky has used the phrase "quantum nondemolition" (QND) to describe a measurement which, *in principle*, can be made time after time on a single system, giving always the same precise result in the absence of external forces (signals). When external forces are present, quantum nondemolition measurements are an ideal tool for monitoring them. Quantum counting can be done accurately and predictably in either a demolition or a nondemolition mode: Photon counting with x-ray proportional counters is demolition; nondemolition methods of counting microwave photons have been proposed by Unruh (1977, 1978) and Braginsky, Vorontsov, and Khalili (1977). Our proposed back-action-evading measurements of the real part of the complex amplitude are nondemolition in principle. For further discussion of the phrase "quantum nondemolition" see Secs. II.E and IV.

B. Technological applications of quantum oscillators coupled to classical forces

The problem of measuring classical signals with a weakly coupled oscillator arises in a variety of contexts—e.g., in experiments to detect gravitational radiation; in the reception of long-wavelength electromagnetic waves using antennas that are very small compared to a wavelength; in experiments to test general relativity (e.g., Eötvös experiments); in gravimeters, gravity gradiometers, accelerometers, and gyroscopic devices (inertial navigators, gyrocompasses, guidance systems); and elsewhere. In most of these areas quantum-mechanical properties of the oscillator are not an issue at present or even in the near future; but they may become an issue in the more distant future—and, equally importantly, the back-action-evading methods of measurement described in this paper may improve the signal-to-noise ratio even in the classical regime.

The task of detecting gravitational waves (Thorne, 1980; introductory review in this issue of *Reviews of Modern Physics*) was the immediate motivation for our interest in quantum-mechanical oscillators as detectors of classical signals. A long-range goal is to detect millisecond-duration bursts of gravitational waves from supernovae at a sufficient distance (the Virgo cluster of galaxies) to guarantee several events per year [see, e.g., Thorne (1978) or Epstein and Clark (1979)]. Bursts from that distance are predicted to have a quantum-mechanical occupation number $n \sim 10^{75}$ for states with the wave vector inside the solid angle, $\Delta\Omega \sim 10^{-38}$ sr, subtended at Earth by the source [cf. Eqs. (6)–(8) of Thorne *et al.* (1979)]. The occupation number averaged over all states in the roughly 45-deg beamwidth of the antenna is $\bar{n} \sim 10^{37}$. This is also the mean number of gravitons that interacts with the antenna during one cycle as the wave burst passes. Clearly, the force of these gravitons on the antenna should be highly classical. Unfortunately, a resonant-bar antenna of mass m couples so weakly to these waves that they can change the number N of phonons in its fundamental mode by only $\delta N \lesssim 0.4\,(N+\frac{1}{2})^{1/2}(\mathrm{m}/10\ \mathrm{tons})$ (cf. Thorne, 1978)—a change so small that with standard "amplitude-and-phase" methods of measurement, the uncertainty principle prevents detection. The 1979 gravitational-wave detectors will be several orders of magnitude away from the amplitude-and-phase limit $\Delta N = (N+\frac{1}{4})^{1/2}$, but the limit might be reached within 5 years.

A Russian experimenter, Vladimir Braginsky, called attention to this problem in 1974 in a series of lectures at American centers for experimental relativity (Stanford, LSU, MIT, Princeton, and Caltech; see Braginsky, 1977). Braginsky and Vorontsov (1974) proposed circumventing the problem by replacing amplitude-and-phase methods with "phonon counting." It did not, and does not, seem practical to count the phonons directly. Instead one might, as Braginsky and Vorontsov suggested, couple the bar to a microwave cavity, thereby converting phonons into microwave photons; measure the number of microwave photons; and thereby monitor changes in the number of phonons in the bar. Braginsky and Vorontsov (1974) proposed a specific method of measuring the number of microwave photons; see also Braginsky, Vorontsov, and Krivchenkov (1975). Three years later Unruh (1977, 1978) proved that this Braginsky–Vorontsov method is flawed, and Braginsky, Vorontsov, and Khalili (1977) found the flaw in their original, unpublished analysis. However, in these same papers Unruh (1977, 1978) and Braginsky *et al.* (1977) proposed new "quantum-nondemolition" (QND) methods of measuring the number of microwave photons—methods that still look viable in principle.

Unfortunately, any QND quantum-counting technique at acoustical or microwave frequencies, including the new Unruh and Braginsky techniques, may be extremely difficult to implement in practice. This is because, to avoid perturbing the number of quanta $\hat{N}$ while measuring it, one must construct an interaction Hamiltonian that commutes with $\hat{N}$; such a Hamiltonian must be quadratic (or quartic or ...) in the amplitude of the oscillator; and at these frequencies it is extremely difficult to construct quadratic couplings with negligible linear admixtures.

In response to this dilemma the authors (Thorne *et al.*, 1978, 1979) proposed using linearly coupled, "back-action-evading" measurements of the real part X_1 of the oscillator's complex amplitude. Such measurements can be performed by modest modifications of the amplitude-and-phase electronic techniques now in use [Eq. (6) of Thorne *et al.* (1978); Sec. 7.2 of Thorne *et al.* (1979); next-to-last section of Braginsky *et al.* (1980); Sec. II.F.3 of this paper]. It now (August 1979) is widely assumed by gravity experimenters that third-generation bar antennas will incorporate back-action-evading electronic techniques.

C. Detailed summary of this paper

This paper serves two purposes: First, it reviews those aspects of the measurement of classical signals with a quantum-mechanical oscillator which we think are important (i) for a conceptual understanding of the subject, and (ii) for the future development of the subject. Second, it presents in detail our own new ideas on back-action-evading measurements of oscillators. This paper does *not* attempt a review of efforts to detect gravitational waves. For that topic in brief see the companion paper (Thorne, 1980); for greater detail see Tyson and Giffard (1978), Braginsky and Rudenko (1978), Douglass and Braginsky (1979), Weber (1979),

or Weiss (1979).

This is Paper I of a two-paper treatise. Paper I deals with issues of principle; and it takes the viewpoint of a theoretical physicist, who enjoys proving theorems by *Gedanken* experiments, and who believes firmly in nonrelativistic quantum theory—including the reduction of the wave function when measurements are made. Paper II will be published in a future issue of *Reviews of Modern Physics*. It will deal with practical realizations of back-action-evading measurements; and its viewpoint will be more nearly that of an experimental physicist or electrical engineer. The two papers will rely little on each other. It should be easy to read one without reading the other, but it may not be easy to wade through either one.

Readers who are awaiting the publication of Paper II may find Thorne *et al.* (1979) and the last few sections of Braginsky, Vorontsov, and Thorne (1980) useful. These two references describe briefly some of the material that Paper II will cover in greater detail.

The body of this paper consists of three major sections: II, III, and IV.

1. Summary of Sec. II

In Sec. II we discuss measurements of a quantum-mechanical oscillator from a somewhat formal mathematical viewpoint.

Section II.A gives examples of the types of oscillators, both mechanical and electromagnetic, that we consider; it explains our neglect of fluctuations due to Nyquist forces (internal friction); and it introduces a single, unified mathematical description of the various oscillators. The most important items in this description are the oscillator's mass m, frequency ω, position $\hat{x}$, momentum $\hat{p}$, number of quanta $\hat{N}$, phase $\hat{\psi}$, and complex amplitude $\hat{X}_1+i\hat{X}_2$, which are related by

$$\hat{x}+i\hat{p}/m\omega=(\hat{X}_1+i\hat{X}_2)e^{-i\omega t},$$
$$\hat{N}=(m\omega/2\hbar)(\hat{X}_1^2+\hat{X}_2^2)-\tfrac{1}{2},\quad \hat{\psi}=\tan^{-1}(\hat{X}_2/\hat{X}_1). \tag{1.1}$$

[Carets ("^") denote quantum-mechanical operators.]

Section II.B first considers the Heisenberg uncertainty principle, which in three equivalent guises states that

$$\left.\begin{array}{ll}\text{Heisenberg}\\ \text{uncertainty}\\ \text{principle}\end{array}\right\}\quad \begin{array}{l}\Delta x\cdot\Delta p\geq\tfrac{1}{2}\hbar,\\ \Delta X_1\cdot\Delta X_2\geq\hbar/2m\omega,\\ \Delta N\cdot\Delta\psi\geq\tfrac{1}{2}.\end{array} \tag{1.2}$$

Here $2\pi\hbar$ is Planck's constant. A simple consequence of the uncertainty principle is that the "error box" in the $(x,p/m\omega)$ [or (X_1,X_2)] phase plane, associated with any measurement of the oscillator, has a minimum area $\pi\hbar/2m\omega$ (Fig. 1 below). Section II.B then introduces three types of measurement. The first type is "*amplitude-and-phase measurement*," for which the error box is round [Fig. 3(b) below]:

$$\left.\begin{array}{ll}\text{standard}\\ \text{quantum}\\ \text{limits}\end{array}\right\}\quad \begin{array}{l}\Delta X_1=\Delta X_2\geq(\hbar/2m\omega)^{1/2},\\ \\ \Delta N\geq(N+\tfrac{1}{4})^{1/2},\quad \Delta\psi\geq\tfrac{1}{2}N^{-1/2}\ \text{for } N\gg1.\end{array} \tag{1.3}$$

Its minimum errors (1.3) are called the "standard quantum limits" because it is the type of measurement made by standard electronic techniques. The second type is "*quantum counting*" (measurement of $\hat{N}$). Since the outcome of a quantum-counting measurement is always an integer N and the measurement leaves the phase ψ completely indeterminate, one can think of quantum counting as having an annular error box [Fig. 3(c) below], which encompasses at least the region between $N-\frac{1}{2}$ quanta and $N+\frac{1}{2}$ quanta:

$$\Delta(X_1^2+X_2^2)^{1/2}\gtrsim N^{-1/2}(\hbar/2m\omega)^{1/2}\ \text{for } N\gg1,$$
$$\psi\ \text{completely indeterminate}. \tag{1.4}$$

Note that if the oscillator is highly excited, $N\gg1$, quantum counting can determine the absolute value of its complex amplitude, $|X|=(X_1^2+X_2^2)^{1/2}$, far more accurately than the standard quantum limit. The third type of measurement is "*back-action-evading measurement* of $\hat{X}_1$," for which the error box is a long, thin ellipse [Fig. 3(d) below], with

$$\Delta X_1\ll\Delta X_2,$$
$$\Delta X_1\ \text{as small as one wishes, in principle}. \tag{1.5}$$

If the experimenter prefers, he can make a back-action-evading measurement of $\hat{X}_2$.

These three types of measurement are analyzed, each in turn, in Secs. II.C, II.D, and II.E with emphasis on the accuracy with which each type of measurement can monitor a classical force $F(t)$. The action of the force between time t_0 and $t_0+\tau$ is characterized by the dimensionless force integral

$$\alpha=(2m\omega\hbar)^{-1/2}\int_{t_0}^{t_0+\tau}F(t')i\exp[i\omega(t'-t_0)]\,dt', \tag{1.6a}$$

which is simply related to the change of complex amplitude that the force would produce if the oscillator were classical:

$$\delta(X_1+iX_2)=(2\hbar/m\omega)^{1/2}\alpha e^{i\omega t_0}. \tag{1.6b}$$

For *amplitude-and-phase measurements* (Sec. II.C) the force is detectable if and only if

$$\text{standard quantum limit: } |\alpha|\gtrsim1; \tag{1.7}$$

cf. Eqs. (1.6b) and (1.3).

For *quantum-counting measurements* (Sec. II.D) the force is detectable if and only if

$$|\alpha|\gtrsim(N+1)^{-1/2}, \tag{1.8}$$

where N is the number of quanta in the oscillator; cf. Eqs. (1.6b) and (1.4). Note that if the oscillator initially is arbitrarily highly excited ($N\to\infty$), then arbitrarily weak forces can be detected by quantum counting. However, quantum counting measurements can never measure the details of the force (can never determine the precise value of $|\alpha|$) with a precision better than a factor ~3—unless the force is so strong that it increases the energy by an amount large compared to the initial energy. This is because in an N-quantum state the initial phase ψ of the oscillator is completely indeterminate; so one cannot know whether the force was acting in such a way as to change predominantly the oscillator's amplitude (number of quanta), or acting in such a way as to change predominantly its phase.

For *back-action-evading measurements* of $\hat{X}_1$ (Sec. II.E), the force can be detected and measured with a precision

$$\Delta[\mathrm{Re}(\alpha e^{i\omega t_0})] = \Delta\left((2m\omega\hbar)^{-1/2} \times \int_{t_0}^{t_0+\tau} F(t')\sin\omega t'\,dt'\right) \ll 1, \quad (1.9)$$

and, in principle, the measurement can be arbitrarily accurate, $\Delta[\mathrm{Re}(\alpha e^{i\omega t_0})] \to 0$, for any measurement time τ. By a sequence of arbitrarily quick and accurate measurements of $\hat{X}_1$, one can monitor the details of $F(t)$ with arbitrary accuracy—except at times t_0 near 0, $\pi/\omega, 2\pi/\omega, \ldots$ when $\sin\omega t_0 \simeq 0$ [cf. Eq. (1.9)]. To get good accuracy near these times one can couple the force to a second oscillator on which one measures $\hat{X}_2$ rather than $\hat{X}_1$, thereby monitoring $\mathrm{Im}(\alpha e^{i\omega t_0})$ with a precision

$$\Delta[\mathrm{Im}(\alpha e^{i\omega t_0})] = \Delta\left((2m\omega\hbar)^{-1/2} \times \int_{t_0}^{t_0+\tau} F(t')\cos\omega t'\,dt'\right) \ll 1. \quad (1.10)$$

Having argued that nonrelativistic quantum theory permits back-action evasion to monitor $\hat{X}_1$ and the classical force $F(t)$ with arbitrary accuracy, Sec. II.E then notes two limits of principle: First, relativistic effects prevent $\hat{X}_1$ from being measured more accurately than the oscillator's Compton wavelength:

$$\Delta X_1 \gtrsim 2\pi\hbar/mc \simeq 2\times 10^{-43}\ \mathrm{cm\ for}\ m = 1\,\mathrm{ton} \quad (1.11)$$

(a limit that is completely irrelevant for macroscopic systems). Second, at some level of accuracy one will discover that the force $F(t)$ is not classical, but rather is carried by discrete bosons (gravitons if F is due to gravitational waves). At this level one's measurements are sensitive to vacuum fluctuations in the system that produces $F(t)$. For measurements that last a time $\tau \gtrsim \omega^{-1}$ this "real quantum limit" is

$$\left.\begin{array}{l}\text{real}\\ \text{quantum}\\ \text{limits}\end{array}\right\} \begin{array}{l}\Delta X_1 \gtrsim (\hbar/2m\omega)^{1/2}(\omega\tau/\bar{n}_{\mathrm{SQL}})^{1/2},\\ \\ |\alpha| \gtrsim (\omega\tau/\bar{n}_{\mathrm{SQL}})^{1/2}.\end{array} \quad (1.12)$$

Here $\bar{n}_{\mathrm{SQL}}$ is the mean occupation number in the quantum-mechanical states associated with F, when F is just barely strong enough to be detected in one cycle by amplitude-and-phase techniques. In the case of today's resonant-bar gravitational-wave antennas $\bar{n}_{\mathrm{SQL}} \simeq 10^{38}$ (cf. Sec. I.B), so the real quantum limit is a factor 10^{19} smaller than the standard quantum limit—so small as to be ridiculously irrelevant in the twentieth century.

Any actual measurement is carried out by coupling the oscillator to an external measuring apparatus. The details of that coupling are embodied, mathematically, in an "interaction Hamiltonian." Section II.F describes three types of back-action-evading measurements of $\hat{X}_1$ and writes down their interaction Hamiltonians. The first type, a *continuous two-transducer measurement*, requires both a position transducer and a momentum transducer. This type of measurement can measure $\hat{X}_1$ arbitrarily quickly and arbitrarily accurately, in principle; and it can lead to an arbitrarily accurate monitoring of a classical force. The second type, a *stroboscopic measurement*, uses a single transducer (position or momentum) with stroboscopically pulsed coupling to the oscillator. For nonzero pulse durations Δt, the back-action evasion is imperfect, and the measurement precision is limited to

$$\Delta X_1 \gtrsim (\hbar/m\omega)^{1/2}(\omega\Delta t)^{1/2}. \quad (1.13)$$

The third type, a *continuous single-transducer measurement*, uses a single transducer with sinusoidally modulated coupling to the oscillator, followed by a filter which averages the transducer output over a time $\tilde{\tau} \gg 2\pi/\omega$. For noninfinite averaging times $\tilde{\tau}$ the back-action evasion is imperfect, and the precision is limited to

$$\Delta X_1 \gtrsim (\hbar/2m\omega)^{1/2}(\omega\tilde{\tau})^{-1/2}. \quad (1.14)$$

Section II.G discusses back-action-evading measurements for zero-frequency ($\omega = 0$) oscillators—i.e., for "free masses." In this case the quantities analogous to $\hat{X}_1$ and $\hat{X}_2$ are $\hat{x} - (\hat{p}/m)t$ and $\hat{p}$ (where $\hat{x}$ is position and $\hat{p}$ is momentum). By back-action-evading measurements of either of these quantities, one can monitor an external force with arbitrary accuracy—in principle.

2. Summary of Sec. III and Appendixes A-D

All of the above limits on measurement accuracy [Eqs. (1.2)–(1.14)] are derived, in Sec. II, from a rather formal viewpoint. This viewpoint pays no attention to the details of the measurement method. Instead it assumes—in line with Neils Bohr's interpretation of nonrelativistic quantum theory—that in an instantaneous measurement any observable, by itself, can be measured arbitrarily accurately, leaving the oscillator afterward in an eigenstate of that observable with eigenvalue equal to the measured result (see footnote 6). This viewpoint assumes, further, that any two observables $\hat{A}$ and $\hat{B}$ can be measured simultaneously with precisions constrained only by the Heisenberg uncertainty principle

$$\Delta A\,\Delta B \geq \tfrac{1}{2}|\langle[\hat{A},\hat{B}]\rangle|. \quad (1.15)$$

Moreover, it assumes that such a measurement can leave the oscillator in a state with expectation values $\langle\hat{A}\rangle$ and $\langle\hat{B}\rangle$ equal to the measured values to within precisions ΔA and ΔB, and with variances of order ΔA and ΔB.

Because Bohr's viewpoint is not universally accepted (see footnote 6 below), some physicists have worried whether the measurement limits derived from it [Eqs. (1.2)–(1.15)] can actually be achieved in principle. Are there, perhaps, other more stringent limits which show up only in a more detailed analysis of the measurement process? Section III and Appendixes A–D prove, for back-action-evading measurements, that more stringent limits do not exist; the limits described above are actually achievable in principle. The method of proof is to present *Gedanken* experiments that actually achieve those limits. The *Gedanken* experiments are sketched in Sec. III; the rather complicated details of

the experimental apparatus are presented in Appendixes A and B; and detailed mathematical analyses of the experiments are given in Appendixes C and D.

It is crucial that we be able to analyze our *Gedanken* experiments fully and exactly using the mathematical techniques of quantum theory. This means, unfortunately, that their details must differ from realistic experiments, with real amplifiers and real electronic readout systems. Realistic experiments will be analyzed using semiclassical techniques in Paper II [see also Thorne *et al.* (1979) and Braginsky *et al.* (1980)].

A crucial feature in our full quantum analysis of the *Gedanken* experiments is the "reduction of the wave function" at the end of each measurement. Most quantum mechanics textbooks talk about this reduction, but they do not present examples or exercises in which the reduction occurs. Therefore, the reader may find interesting in itself our use of the reduction of the wave function in Appendix C. There we analyze a sequence of measurements of $\hat{X}_1$, and the reduction of the wave function allows us to carry the quantum-mechanical analysis from one measurement to the next. Of similar interest may be our mathematical model for an imprecise "readout system" in Sec. 7 of Appendix C.

A key element in our *Gedanken* experiments is the coupling of the oscillator to the measuring apparatus. In practice that coupling occurs in a transducer; mathematically it is embodied in the interaction Hamiltonian $\hat{H}_I$; in either case the strength of the coupling can be described by a dimensionless coupling constant

$$\beta \simeq \frac{\text{energy stored in measuring system due to transducer coupling}}{\text{energy of oscillator's motions}} . \tag{1.16}$$

To achieve the limiting precisions of Sec. II [Eqs. (1.2)–(1.15)], one must use very strong coupling: $\beta \gtrsim 1$ typically; $\beta \to \infty$ in some cases. When β is fixed at some modest value by practical considerations, the limits of Sec. II get replaced by new "weak coupling measurement limits." These new limits are revealed as by-products of the *Gedanken* experiments of Sec. III and Appendixes A–D. Alternative derivations of the weak coupling limits will be given in the sequel to this paper (Paper II), using realistic models for the measurements and using electrical-engineering techniques of analysis. Some of our alternative derivations are also sketched by Braginsky, Vorontsov, and Thorne (1980).

Turn now to a blow-by-blow overview of Sec. III and Appendixes A–D.

Section III.A analyzes measurements of a free mass (oscillator with $\omega = 0$). As a prelude to the analysis, Sec. III.A.1 derives a standard quantum limit for the monitoring of a constant force F acting on a free mass:

$$\text{standard quantum limit: } \Delta F \gtrsim (m\hbar/\tau^3)^{1/2} . \tag{1.17}$$

Here m is the mass and τ is the duration of the measurement. This is the analog of the standard quantum limit for amplitude-and-phase measurements of an oscillator [Eq. (1.7)]. It is derived in Sec. III.A for two types of free-mass measurements: measurements of position $\hat{x}$ and measurements of velocity. By contrast, back-action-evading measurements of momentum $\hat{p}$ (analogue of $\hat{X}_2$) can be arbitrarily accurate in principle

$$\Delta F \to 0, \quad \tau \to 0 . \tag{1.18}$$

This is proved by a *Gedanken* experiment in Sec. III.A.2. A by-product of that proof is the limit

$$\Delta F \gtrsim (m\hbar/\tau^3)^{1/2}\beta^{-1/2} \tag{1.19}$$

[Eq. (3.15)] for back-action-evading measurements of $\hat{p}$ with finite coupling constant β. Back-action-evading measurements of $\hat{p}$ require a momentum transducer. Section III.A.2 shows that a momentum transducer is equivalent to a velocity transducer plus a capacitor with negative capacitance. Velocity transducers are easy to construct, but negative capacitors are not. Appendix A presents models for several types of negative capacitors and shows that one type—the "spring-based negative capacitor" (Appendix A.1)—can perform the required role in a momentum transducer, without introducing any noise into the back-action-evading measurement of $\hat{p}$ (Appendix A.2). For the reader who feels uneasy about negative capacitors, Appendix A.3 describes several alternative viewpoints that may allay his uneasiness. The reader who is still unhappy perhaps can find some comfort in Appendix A.5, which shows that in slightly modified *Gedanken* experiments an inductor can do an adequate job as a negative capacitor.

Section III.B analyzes continuous two-transducer back-action-evading measurements of an oscillator. A *Gedanken* experiment is sketched, which proves that $\hat{X}_1$ can be measured arbitrarily quickly and accurately; this, in turn, allows an external force to be monitored with perfect precision. Section III.B gives the Hamiltonian for the *Gedanken* experiment [Eq. (3.16)], and Appendix B shows how that Hamiltonian could be realized with physical measuring apparatus for the case of a mechanical oscillator (Appendix B.1) and for an electromagnetic oscillator (Appendix B.2). The physical realization requires a noise-free negative capacitor (Appendix A) for the mechanical case, and a noise-free negative spring (Fig. 11) for the electromagnetic case. A mathematical analysis of the *Gedanken* experiment is sketched in Sec. III.B, and is presented in detail—including reduction of the wave function after each measurement—in Appendix C. A by-product of the *Gedanken* experiment is the following limit on the precision of a continuous two-transducer back-action-evading measurement of $\hat{X}_1$ when the coupling constant β is finite [Eq. (3.21)]:

$$\Delta X_1 \gtrsim (\hbar/2m\omega)^{1/2}(\beta\omega\tau)^{-1/2} . \tag{1.20}$$

Here τ is the duration of the measurement.

It is not likely that in real experiments one can construct negative capacitors or negative springs with the low-noise performance required by momentum trans-

ducers. Consequently, experimenters may be forced to perform back-action-evading measurements with only a position transducer. Such measurements can be of two types: stroboscopic (pulse duration $\Delta t \ll 1$), or continuous (filter averaging time $\tilde{\tau} \gg 2\pi/\omega$); see above.

Appendix D presents a *Gedanken*-experiment analysis of continuous single-transducer measurements of X_1. That analysis distinguishes two different averaging times: (i) the averaging time $\tilde{\tau}$ of the filter, which follows the modulated transducer and (in a realistic experiment) precedes the amplifier; (ii) the total time τ over which the experimenter averages the filter's output, in order to arrive at a final result for X_1. In realistic experiments the total averaging time τ is greater than or equal to the filter averaging time $\tilde{\tau}$. The filter partially protects $\hat{X}_1$ from back action. Its failure to give complete protection leads to an absolute limit

$$\Delta X_1 \gtrsim (\hbar/2m\omega)^{1/2}(\omega\tilde{\tau})^{-1/2} \tag{1.21a}$$

on the precision of the measurement [Eq. (1.14); Eq. (D21)]. When the coupling is weak, $\beta \lesssim \tilde{\tau}/\tau \leq 1$, Appendix D reveals a more stringent limit

$$\Delta X_1 \gtrsim (\hbar/2m\omega)^{1/2}(\beta\omega\tau)^{-1/2} \tag{1.21b}$$

[Eq. (D19)]. For fixed β, the longer the experimenter averages his signal, the greater will be his accuracy ($\Delta X_1 \propto \tau^{-1/2}$)—until he hits the "accuracy floor" (1.21a) determined by the averaging time $\tilde{\tau}$ of his filter.

This paper does not give an explicit *Gedanken*-experiment analysis of stroboscopic measurements. However, the limiting stroboscopic precisions can be read off other analyses given in the paper. The limiting precisions involve two time scales: (i) the duration of each stroboscopic measurement (the "pulse time") $\Delta t \ll \pi/\omega$; and (ii) the total time τ over which the experimenter averages his measurements—or, equivalently, the total number $\mathfrak{N} \equiv \omega\tau/\pi$ of measurements that he averages to get a single value for X_1. The finite pulse time Δt leads to imperfect evasion of back action and an absolute limit

$$\Delta X_1 \gtrsim (\hbar/m\omega)^{1/2}(\omega\Delta t)^{1/2} \tag{1.22a}$$

on the measurement precision [Eq. (1.13); Sec. II.F.2]. When the coupling is weak, $\beta \lesssim (\mathfrak{N}^{1/2}\omega\Delta t)^{-2}$, there is a more stringent limit. This limit can be read off that for measurements of free masses, since the oscillator behaves essentially like a free mass during each pulsed measurement ($\Delta t \ll \pi/\omega$). Specifically, the free-mass weak coupling limit (3.15a), with $\tau \to \Delta t$ and $\delta p_0 \to m\omega\Delta X_1$, and with an added factor $\mathfrak{N}^{-1/2}$ to account for averaging of $\mathfrak{N}$ data points, becomes

$$\Delta X_1 \gtrsim (\hbar/m\omega)^{1/2}(\mathfrak{N}\beta\omega\Delta t)^{-1/2}. \tag{1.22b}$$

(This limit will be derived more carefully in Paper II.) For fixed β and Δt this weak coupling limit improves as $\mathfrak{N}^{-1/2}$ until one hits the "accuracy floor" (1.22a). Note that there is an optimal pulse time Δt which leads to an absolute minimum for the sum of the errors (1.22a) and (1.22b):

$$(\Delta t)_{\text{optimum}} \simeq \omega^{-1}(\beta\mathfrak{N})^{-1/2}, \qquad \Delta X_1 \gtrsim (\hbar/m\omega)^{1/2}(\beta\mathfrak{N})^{-1/4}. \tag{1.22c}$$

This absolute minimum stroboscopic limit improves only as $\mathfrak{N}^{-1/4}$, when one combines the results of many measurements. By contrast the limits (1.21) for continuous single-transducer back-action-evading measurements improve as (averaging time)$^{-1/2}$. The continuous measurements are better because of their larger duty cycle.

In real experiments one must face not only the above quantum limits, but also limits due to (i) noise from a real amplifier in the readout circuit, and (ii) noise from Nyquist (frictional) forces in the oscillator. These will be discussed in Paper II. Here we summarize the results: Amplifier noise modifies all of the above limits by the simple replacement

$$\hbar \to 2kT_n/\Omega, \tag{1.23}$$

where T_n is the "noise temperature" of the amplifier and Ω is the frequency at which it operates (Braginsky, Vorontsov, and Khalili, 1978; Thorne *et al.*, 1979; Braginsky *et al.*, 1980). Nyquist noise produces a change in X_1, during time τ, given by

$$\Delta X_1 \simeq (kT/m\omega^2)^{1/2}(\omega\tau/Q)^{1/2}. \tag{1.24}$$

Here T is the oscillator's physical temperature and Q is its "quality factor."

Equations (1.20)–(1.24) are a complete set of limits on the accuracy of realistic back-action-evading measurements of oscillators. However, to apply them in a specific case one must be able to evaluate the coupling constant β in terms of the actual experimental parameters. This issue will be discussed in Paper II; see also Braginsky, Vorontsov, and Thorne (1980).

3. Summary of Sec. IV

Section IV develops a formal mathematical theory of quantum-nondemolition measurements—a theory which generalizes to arbitrary quantum-mechanical systems the oscillator and free-mass results of Secs. II and III.

There are two different, but closely related viewpoints on quantum nondemolition measurement.

The first viewpoint focuses on the uncertainties which are built into a quantum-mechanical description of the measured system, and it ignores the details of the system's coupling to a measuring apparatus. This viewpoint is most easily introduced, perhaps, by considering a particularly simple system—a free mass. For a free mass there is a sharp difference between position $\hat{x}$ and momentum $\hat{p}$. Either can be measured arbitrarily accurately in an instantaneous measurement. However, an initial precise measurement of $\hat{x}$ perturbs $\hat{p}$ strongly ($\Delta p \geq \hbar/2\Delta x$); during subsequent free evolution, $\hat{p}$ drives changes in $\hat{x}$ [$\hat{x}(t) = \hat{x}(0) + \hat{p}(0)t/m$]; and, as a result, the accuracy of a second measurement of $\hat{x}$ is spoiled:

$$[\Delta x(t)]^2 = [\Delta x(0)]^2 + [\Delta p(0)/m]^2 t^2 \geq [\Delta x(0)]^2 + [\hbar/2m\Delta x(0)]^2 t^2 \geq \hbar t/m. \tag{1.25}$$

By contrast, an initial precise measurement of $\hat{p}$ perturbs $\hat{x}$ strongly; but since $\hat{x}$ does *not* feed into $\hat{p}$ during free evolution, a second measurement of $\hat{p}$ at a later time can be just as accurate as the first—arbitrarily accurate, in fact. The first viewpoint generalizes to arbitrary systems this special property of $\hat{p}$. It characterizes a "quantum nondemolition (QND) observable" as an observable which, like $\hat{p}$, can be measured over and over again with arbitrarily great accuracy. The essential feature of a QND observable is that its free evolution, like that of $\hat{p}$, is isolated from observables with which it does not commute. For an oscillator the QND observables include $\hat{X}_1$, $\hat{X}_2$, and the number of quanta $\hat{N}$.

Section IV.A adopts this first viewpoint. It defines precisely the concept of a QND observable; it derives from the definition several important formal properties of QND observables; and it delineates various types of QND observables: stroboscopic QND observables, continuous QND observables, generalized QND observables, and QNDF observables. The last of these are "QND" in the presence of a classical force as well as in its absence; they can be used, in principle, to monitor the details of the force with arbitrary accuracy. An example is $\hat{X}_1$. Observables which are QND but not QNDF can be used to detect the presence of an arbitrarily weak force, but they cannot determine its precise strength. An example is the number of quanta $\hat{N}$ in an oscillator .

The second viewpoint on quantum nondemolition measurement focuses on the quantum-mechanical nature of the measuring apparatus. It characterizes a QND observable as one that can be completely shielded from the back action of the measuring apparatus. Any observable one chooses to measure can be free of *direct* back action from the measuring apparatus, provided the interaction between the system and the measuring apparatus is designed properly (the measured observable being the only observable of the system which appears in the interaction Hamiltonian). However, for most observables the measuring apparatus will act back indirectly through observables which do not commute with the measured observable. For example, in a measurement of free-mass position $\hat{x}$, the measuring apparatus acts back on $\hat{p}$, which then drives changes in $\hat{x}$ via free evolution. Clearly, this second viewpoint is closely related to the first; the essential feature of both is that the evolution of a QND observable is isolated from observables with which it does not commute. Section IV.B shows that the two viewpoints are essentially equivalent by proving that any QND observable (defined using the first viewpoint) can be completely shielded from the back action of the measuring apparatus. The second viewpoint has been used by Unruh (1979) to characterize quantum nondemolition measurement.

Section IV.C, which concludes this paper, discusses the application of the general QND theory of Secs. IV.A and IV.B to real experiments. A warning is given that the general theory is too simplistic. That theory gives sufficient conditions for an experiment to have high precision, but not necessary conditions. By being clever one might be able to violate that theory's conditions and still achieve high accuracy.

II. FORMAL DISCUSSION OF MEASUREMENTS OF HARMONIC OSCILLATORS

A. Mathematical description of the oscillator

The oscillators that we study are macroscopic in size. An example is the fundamental mode of mechanical oscillation of a monocrystal of sapphire with mass $M \sim 100$ kg. Such crystals, cooled to a few millidegrees, might be used 5 to 10 years hence as third-generation resonant-bar detectors for gravitational waves [cf. Braginsky (1974) and the lectures by Braginsky, Douglass, and Weber in Bertotti (1977)]. Such a crystal contains $\sim 3\times 10^{27}$ atoms, and therefore its mechanical oscillations have $\sim 3\times(3\times 10^{27})$ degrees of freedom. The fundamental mode is one of those degrees of freedom, and it is almost completely decoupled from all the others. The strength of its coupling to other modes is quantified by its "Q"—which is the number of radians of oscillation required for its energy to decrease by a factor $1/e$ (due to "friction" against the other modes), from an initial energy far above thermal. A Q of 4.2×10^9 has been achieved with a small doubly convex quartz lens at 2 °K by Smagin (1974); a Q of 5×10^9 has been achieved for a 1 kg sapphire crystal at 4.3 °K by Bagdasarov *et al.* (1977) [see also lecture by Braginsky in Bertotti (1977)]; a Q of 2×10^9 has been achieved with a 4.9 kg silicon crystal at 3.5 °K by McGuigan *et al.* (1978); and it is not unreasonable to hope for $Q \sim 10^{13}$ at a temperature of a few millidegrees.

The coupling to other modes produces not only friction; it also produces fluctuating forces ("Nyquist forces") which cause the amplitude of the fundamental mode to random walk. In thermal equilibrium the mean number of phonons in the fundamental mode is $\bar{N} = kT/\hbar\omega \sim 10^4$ for $T \sim 0.003$ °K and $\omega/2\pi \sim 5000$ Hz. In a time interval $\Delta t \ll Q/\omega$ the number of phonons random walks by $\Delta N \sim \bar{N}(2\omega\Delta t/Q)^{1/2}$. Hence, a change of unity requires a mean time of $\Delta t \sim Q/(2\omega\bar{N}^2) \sim 1$ sec if $Q \sim 10^{13}$. This is very long compared to the 0.2 msec period of the fundamental mode—so long, in fact, that for such a crystal Nyquist forces should be totally negligible compared to noise and quantum-mechanical uncertainties in the device that measures the crystal's oscillations.

Unruh (1980) has recently analyzed quantum-mechanically the effect of thermal (Nyquist) fluctuations on a harmonic oscillator. In his analysis Unruh considers a simple model of an oscillator coupled to a heat reservoir; the heat reservoir consists of a large number of oscillators, each of which is coupled linearly to the primary oscillator.

An obvious second example of a macroscopic oscillator is an electrical LC circuit.

A third example is a normal mode of electromagnetic oscillation of a microwave cavity with superconducting walls. Such cavities are being used as displacement sensors for resonant-bar gravitational-wave detectors (Braginsky, Panov *et al.*, 1977), and they have been proposed as the fundamental element in a new type of gravitational-wave detector (Braginsky *et al.*, 1973; Grishchuk and Sazhin, 1975; Pegoraro, Picasso, and Radicati, 1978; Caves, 1979) and in other gravity experiments (Braginsky, Caves, and Thorne, 1977). The normal modes of such a cavity have Q's of $\sim 10^{11}$ to 10^{12} (Pfister, 1976;

Allen *et al.*, 1971)—high enough that for some purposes one can ignore thermal (Nyquist) fluctuations in the electromagnetic field.

Nyquist forces not only are negligible in some contexts of interest; they are also irrelevant to the issues of principle which this paper addresses. Therefore we shall ignore them until Paper II—i.e., we shall assume that the one mode of interest can be treated as a harmonic oscillator which couples only to (i) the weak classical signal which we seek to measure, and (ii) our measuring system.

The oscillator is characterized by its canonical coordinate $\hat{x}$ and momentum $\hat{p}$, which are Hermitian operators (observables), and by its mass m and angular frequency ω. If the oscillator is the fundamental mode of a resonant bar, we shall normalize $\hat{x}$ to equal the displacement from equilibrium of the end of the bar. Then m will be roughly half the mass of the bar; and, when the bar is decoupled from the measuring apparatus, $\hat{p}$ will be approximately the momentum of the right half of the bar relative to its center. If the oscillator is an LC circuit, we shall normalize $\hat{x}$ to equal the charge on the capacitor. Then m will be the inductance, and $\hat{p}$ will be the magnetic flux in the inductor. If the oscillator is a normal mode of a microwave cavity, we shall normalize m to equal unity. Then $\hat{x}$ can be $(V/4\pi\omega^2)^{1/2}\times$(mean magnetic field in cavity), and $\hat{p}$ can be $(V/4\pi)^{1/2}\times$(mean electric field in cavity), where V is the cavity volume.

No matter what the nature of the oscillator may be, its coordinate and momentum have the commutator

$$[\hat{x},\hat{p}]=i\hbar; \tag{2.1}$$

its Hamiltonian is

$$\hat{H}_0\equiv\hat{p}^2/2m+\tfrac{1}{2}m\omega^2\hat{x}^2; \tag{2.2}$$

its creation and annihilation operators are

$$\hat{a}^\dagger=(m\omega/2\hbar)^{1/2}(\hat{x}-i\hat{p}/m\omega),$$
$$\hat{a}=(m\omega/2\hbar)^{1/2}(\hat{x}+i\hat{p}/m\omega); \tag{2.3}$$

and the operator representing the number of quanta is

$$\hat{N}=\hat{a}^\dagger\hat{a}=\hat{H}_0/\hbar\omega-\tfrac{1}{2}. \tag{2.4}$$

In addition to these standard operators, which one finds in most quantum mechanics textbooks, it is useful to introduce the quantities

$$\hat{X}_1(\hat{x},\hat{p},t)\equiv\hat{x}\cos\omega t-(\hat{p}/m\omega)\sin\omega t, \tag{2.5a}$$
$$\hat{X}_2(\hat{x},\hat{p},t)\equiv\hat{x}\sin\omega t+(\hat{p}/m\omega)\cos\omega t. \tag{2.5b}$$

It is straightforward to show that

$$\hat{x}+i\hat{p}/m\omega=(\hat{X}_1+i\hat{X}_2)e^{-i\omega t}. \tag{2.6}$$

Thus $\hat{X}_1+i\hat{X}_2$ is the quantum-mechanical analog of the oscillator's classical "complex amplitude." As in the classical limit, so also in the Heisenberg picture of quantum mechanics, $\hat{X}_1$ and $\hat{X}_2$ are conserved in the absence of interactions with the outside world:

$$\frac{d\hat{X}_j}{dt}=\frac{\partial\hat{X}_j}{\partial t}-\frac{i}{\hbar}[\hat{X}_j,\hat{H}_0]=0. \tag{2.7}$$

$\hat{X}_1$ and $\hat{X}_2$ are Hermitian operators and are therefore observables. One can show that they, and linear combinations of them with constant coefficients, are the only conserved observables that are linear functions of $\hat{x}$ and $\hat{p}$. Note that $\hat{X}_1$ and $\hat{X}_2$ have explicit time dependence [Eqs. (2.5)]. In this they differ from all the other observables considered above ($\hat{x},\hat{p},\hat{H}_0,\hat{N}$) and from most, but not all, observables that one encounters in quantum theory.

B. Uncertainty principle and ways to measure the oscillator

In classical theory it is possible to measure the oscillator's complex amplitude $X\equiv X_1+iX_2$ with complete precision. Not so in quantum theory. Equations (2.1) and (2.5) imply that $\hat{X}_1$ and $\hat{X}_2$ do not commute:

$$[\hat{X}_1,\hat{X}_2]=i\hbar/m\omega. \tag{2.8}$$

Therefore the variances of $\hat{X}_1$ and $\hat{X}_2$ in any oscillator state must satisfy

$$\Delta X_1\Delta X_2\geq\tfrac{1}{2}|\langle[\hat{X}_1,\hat{X}_2]\rangle|=\hbar/2m\omega, \tag{2.9a}$$

which is the complex-amplitude analog of the Heisenberg uncertainty principle for position and momentum:

$$\Delta x\Delta p\geq\tfrac{1}{2}\hbar. \tag{2.9b}$$

One can think of x and $p/m\omega$ as Cartesian coordinates in a phase plane (we divide by $m\omega$ to make both coordinates have dimensions of length). Then X_1 and X_2 are Cartesian coordinates that rotate clockwise with angular velocity ω relative to the $(x,p/m\omega)$ coordinates [cf. Eqs. (2.5) and (2.6)]. The uncertainty relations $\Delta X_1\Delta X_2\geq\hbar/2m\omega$, $\Delta x\Delta p/m\omega\geq\hbar/2m\omega$ are equivalent manifestations of the fact that any quantum-mechanical state is characterized by an "error box" in the phase plane with area at least $\pi\hbar/2m\omega$; see Fig. 1.

The standard method for measuring the motion of a macroscopic oscillator is to couple it to a canonical-coordinate (x) transducer whose output is proportional to x, and to feed the output into an amplifier. Figure 2 shows a simple example where the oscillator is an LC circuit (part a, to left of dashed line). In this example, x is the charge on the capacitor, p is the flux through the inductor, no transducer is needed, and the amplifier (part b) produces an output voltage $A\cdot Q$ proportional to the total charge Q that flows through it ("zero-impedance charge amplifier"). The amplifier necessarily is noisy. As a minimum, it has noise due to uncertainty-principle constraints on its internal dynamical variables. If this is its only noise, it is called an "ideal amplifier." Viewed non-quantum-mechanically, the noise is of two types: (i) a stochastically fluctuating noise current $I_n(t)\equiv dQ_n/dt$ which, in the case of Fig. 2, gets superimposed on the amplifier's input [so $V_{\rm out}=A\cdot(x+Q_n)$]; and (ii) a noise voltage $V_n(t)$ which, in Fig. 2, produces a driving force on the oscillator and thereby changes its momentum ($\dot{p}=L\ddot{x}=V_n$).

It is useful to distinguish two types of measurements that can be made with such a system: "quick measurements" and "amplitude-and-phase measurements."

In a *quick measurement* one reads out the amplifier output in a time τ short compared to the oscillator period ($\omega\tau\ll1$; "broad-band amplifier"). From the output one infers the instantaneous coordinate x of the

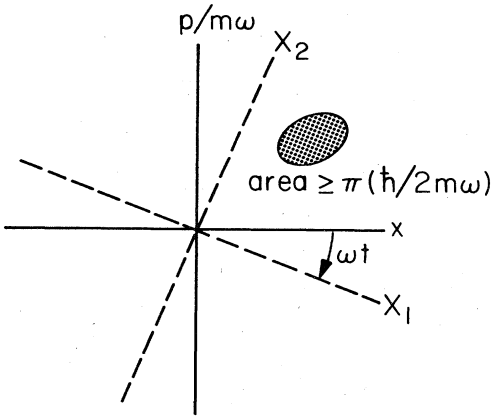

FIG. 1. "Error box" in the phase plane for a quantum-mechanical oscillator. This error box is an ellipse, with centroid at the expectation value $(\langle\hat{x}\rangle, \langle\hat{p}/m\omega\rangle)$ of the position and momentum. The principal axes of the error ellipse are the eigendirections of the variance matrix

$$\begin{Vmatrix} (\Delta x)^2 & \sigma_{xp} \\ \sigma_{xp} & (\Delta p/m\omega)^2 \end{Vmatrix},$$

and the principal radii are the square roots of the corresponding eigenvalues. Here $\sigma_{xp} \equiv (1/2m\omega)\langle(\hat{x}-\langle\hat{x}\rangle)(\hat{p}-\langle\hat{p}\rangle) + (\hat{p}-\langle\hat{p}\rangle)(\hat{x}-\langle\hat{x}\rangle)\rangle$. This error box has the property

$$\Delta x \cdot \frac{\Delta p}{m\omega} \geq \frac{1}{\pi}(\text{area of box}) \geq \frac{\hbar}{2m\omega},$$

$$\Delta X_1 \cdot \Delta X_2 \geq \frac{1}{\pi}(\text{area of box}) \geq \frac{\hbar}{2m\omega}$$

(as one can verify by elementary calculations). Here X_1 and X_2 are the real and imaginary parts of the complex amplitude, and the (X_1, X_2) coordinates of the phase plane are related to the $(x, p/m\omega)$ coordinates by a simple time-dependent rotation

$$x + ip/m\omega = (X_1 + iX_2)e^{-i\omega t}.$$

oscillator to within a precision, for the example in Fig. 2,

$$\delta x_0 \simeq (S_Q/2\tau)^{1/2}. \tag{2.10a}$$

Here S_Q is the spectral density of the amplifier's noise charge $Q_n(t) \equiv \int I_n dt$, and $1/2\tau$ is the bandwidth of the measurement. During this measurement the amplifier's "back-action" noise voltage $V_n(t)$ kicks the oscillator, producing an unpredictable momentum change:

$$\delta p_0 = \int V_n dt \simeq \tfrac{1}{2}(S_V/2\tau)^{1/2}\tau, \tag{2.10b}$$

where S_V is the spectral density of the noise voltage $V_n(t)$.[3] The Heisenberg uncertainty principle places the constraint

$$(S_V S_Q)^{1/2} \geq 2\hbar \tag{2.11}$$

on the noise performance of any zero-impedance charge amplifier [cf. Weber (1959); Heffner (1962); Eq. (3.7) below]. Thus, even with an ideal amplifier, a quick measurement produces an uncertainty product

$$\delta x_0 \cdot \delta p_0 \simeq \tfrac{1}{4}(S_V S_Q)^{1/2} \geq \tfrac{1}{2}\hbar. \tag{2.12}$$

This simple example illustrates how the Heisenberg uncertainty principle is enforced in any quick measurement of precision δx_0: Back-action forces from the measuring system always perturb the oscillator's momentum by an amount $\delta p_0 \gtrsim (\frac{1}{2}\hbar)(1/\delta x_0)$.

A quick measurement produces an uncertainty error box which, for $\delta x_0 \ll \delta p_0/m\omega$, is a long thin ellipse in the phase plane [Fig. 3(a)]. As time passes, the oscillator's "system point" rotates clockwise in the phase plane:

$$x + ip/m\omega = (X_1 + iX_2)e^{-i\omega t}, \quad X_1 + iX_2 = \text{const} \tag{2.13}$$

[Eqs. (2.6) and (2.7)]; and thus its error box also rotates clockwise; see Fig. 3(a). As a result, if one tries to predict the outcome of a second quick measurement of x, the error in the prediction oscillates in time between δx_0 and $\delta p_0/m\omega \geq (\hbar/2m\omega)(1/\delta x_0)$:

$$\delta x(t) = [(\delta x_0)^2 \cos^2\omega t + (\delta p_0/m\omega)^2 \sin^2\omega t]^{1/2}. \tag{2.14}$$

If one wants the maximum of these oscillations to be as small an error as possible, one must arrange for the error box to be round and to have the minimum allowed radius $\delta x_0 = \delta p_0/m\omega = \delta X_1 = \delta X_2 = (\hbar/2m\omega)^{1/2}$. An ideal measurement with these uncertainties will necessarily drive the oscillator into a quantum-mechanical "coherent state"—i.e., a state with a minimum-uncertainty Gaussian wave packet that undergoes classical, oscillatory motion without spreading; see, e.g., Merzbacher (1970).

Turn now from "quick measurements" to "*amplitude-and-phase measurements.*" In such measurements one uses an amplifier that amplifies only a narrow band of frequencies $\Delta\omega \ll \omega$ centered on the oscillator frequency ω. Such an amplifier produces a sinusoidal output with complex amplitude $(V_1 + iV_2) = A \cdot (\bar{X}_1 + i\bar{X}_2)$, where $\bar{X}_1 + i\bar{X}_2$ is a time average of the oscillator's amplitude

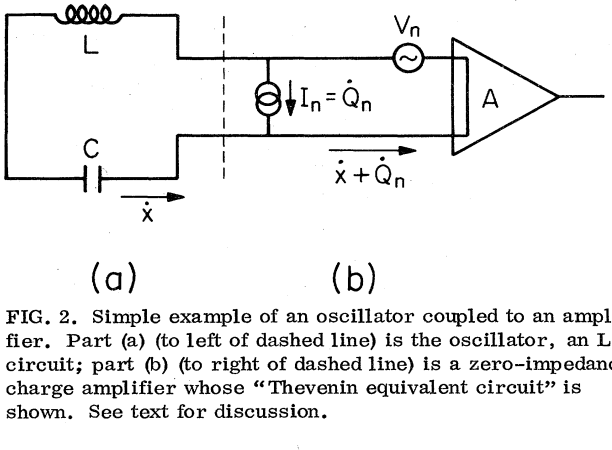

FIG. 2. Simple example of an oscillator coupled to an amplifier. Part (a) (to left of dashed line) is the oscillator, an LC circuit; part (b) (to right of dashed line) is a zero-impedance charge amplifier whose "Thevenin equivalent circuit" is shown. See text for discussion.

[3] A more careful discussion would pay attention to the back-action kick [Eq. (2.10b)] which occurs *during* the initial quick measurement of x. That kick modifies the initial measurement error (2.10a) to read $\delta x_0 \simeq \{S_Q/2\tau + [(\delta p_0/m)\tau]^2\}^{1/2} \simeq (S_Q/2\tau + S_V\tau^3/8m^2)^{1/2}$. The discussion in the text implicitly assumes that the second term is much smaller than the first.

Later (Sec. III.A.1) we shall discuss the case where the two terms are of comparable size. This case leads to an absolute minimum value for the error in our initial quick measurement: $\delta x_0 \gtrsim [(S_Q S_V)^{1/2}\tau/2m]^{1/2} \gtrsim (\hbar\tau/m)^{1/2}$ [cf. Eq. (2.11)].

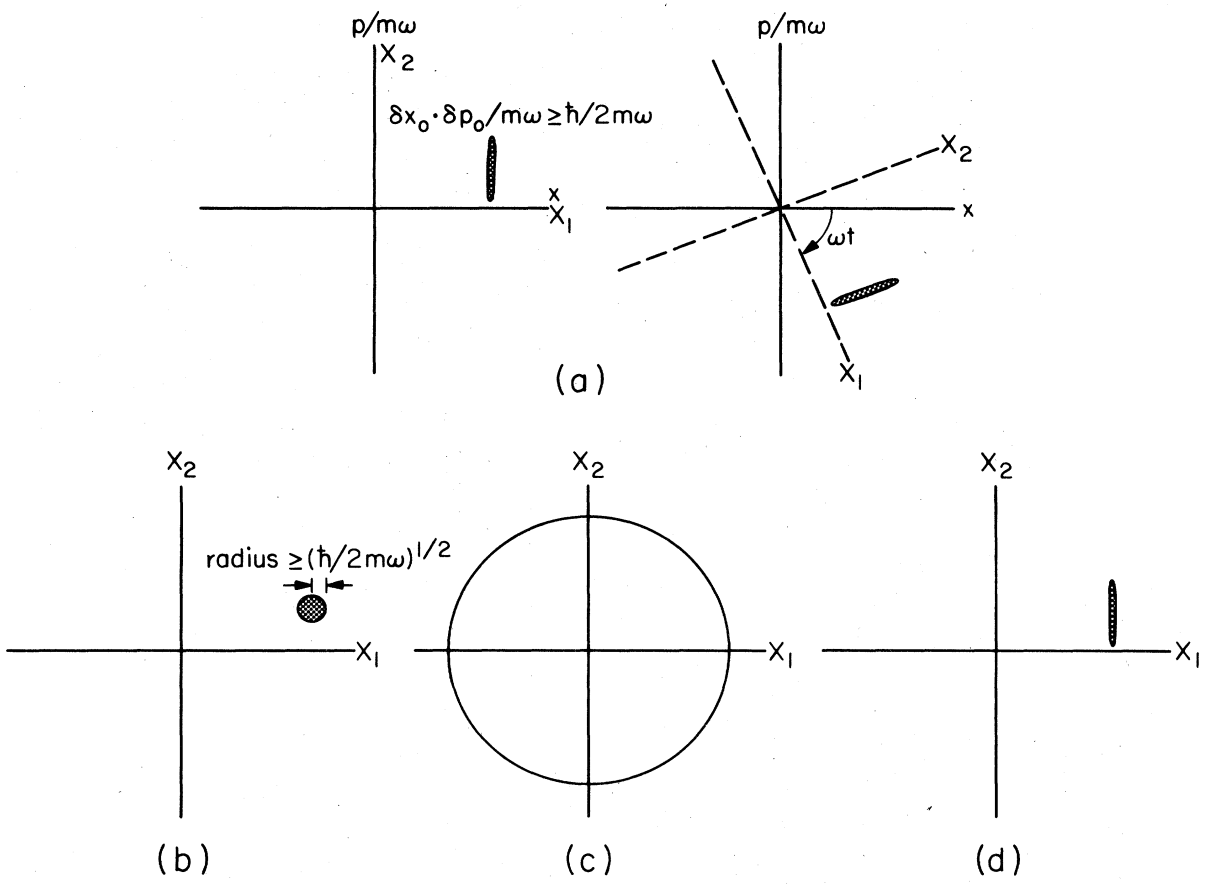

FIG. 3. Error boxes for various types of measurements of a harmonic oscillator. (a) The error box characterizing the results of a "quick measurement" of position. After the measurement the error box rotates clockwise in the phase plane with angular velocity ω, which means that it remains fixed as seen in the "rotating" (X_1, X_2) coordinates. (b) The error box for "amplitude-and-phase measurements" as seen in the (X_1, X_2) coordinate system. (c) The error annulus ($\delta N \equiv 1$) for "quantum-counting measurements." (d) The error box for a "back-action-evading measurement" of X_1.

(averaging time $\tau = \pi/\Delta\omega \gg 1/\omega$). The accuracy of the measurement is constrained by the amplifier's superimposed noise (Q_n in Fig. 2), and by its back-action noise (V_n in Fig. 2). These noises affect the measured amplitudes $\bar{X}_1$ and $\bar{X}_2$ equally (neither phase is preferred), producing the following probable error when S_V/S_Q is optimized:

$$\delta\bar{X}_1 = \delta\bar{X}_2 \simeq [\tfrac{1}{2}(S_V S_Q)^{1/2}/2m\omega]^{1/2} \geq (\hbar/2m\omega)^{1/2} . \qquad (2.15)$$

In the complex-amplitude plane (phase plane) the error box is round; see Fig. 3(b). We call such measurements "amplitude-and-phase" because they attempt to determine both the oscillator's absolute amplitude $|X| \equiv |X_1 + iX_2| = (X_1^2 + X_2^2)^{1/2}$ (or equivalently its energy or number of quanta), and its phase[4] $\psi \equiv \tan^{-1}(X_2/X_1)$.

An "ideal" amplitude-and-phase measurement (one with the minimum possible noise) will drive the oscillator into a quantum-mechanical coherent state with a round error box of radius $\Delta x = \Delta p/m\omega = \Delta X_1 = \Delta X_2 = (\hbar/2m\omega)^{1/2}$. Moreover, for such an ideal measurement the probability distribution of the measured values of X_1 and X_2 is a two-dimensional Gaussian, centered on the expectation value $(\langle\hat{X}_1\rangle, \langle\hat{X}_2\rangle)$ of $\hat{X}_1$ and $\hat{X}_2$ with variances $\Delta X_1 = \Delta X_2 = (\hbar/2m\omega)^{1/2}$. From the measured values of X_1 and X_2, one can infer the oscillator's number of quanta and its phase. It is easy to verify from the Gaussian distribution that the expected value of the inferred number of quanta is $N = (m\omega/2\hbar)(\langle\hat{X}_1\rangle^2 + \langle\hat{X}_2\rangle^2)$, and the variance is $\Delta N = (N + \frac{1}{4})^{1/2}$. For large N the expected value of the inferred phase is $\psi = \tan^{-1}(\langle X_2\rangle / \langle X_1\rangle)$, and the variance is $\Delta\psi = \frac{1}{2}N^{-1/2}$. These variances associated with a coherent state are the minimum possible errors obtainable by the amplitude-and-phase method.

Henceforth, we shall call these minimum errors the "standard quantum limits" for amplitude-and-phase measurements:

$$\left.\begin{array}{l}\text{standard}\\ \text{quantum}\\ \text{limits}\end{array}\right\} \quad \begin{array}{l}\Delta x = \Delta p/m\omega = \Delta X_1 = \Delta X_2 = (\hbar/2m\omega)^{1/2},\\ \Delta N = (N + \frac{1}{4})^{1/2}, \quad \Delta\psi = \frac{1}{2}N^{-1/2} \quad \text{for } N \gg 1 .\end{array} \qquad (2.16)$$

The fact that these are the very best measurement precisions achievable by the amplitude-and-phase method

[4]For a discussion of difficulties with making rigorous the quantum-mechanical concept of the oscillator's phase ψ see, e.g., Carruthers and Nieto (1965). We circumvent these difficulties by working with the real and imaginary parts of the complex amplitude, $\hat{X}_1$ and $\hat{X}_2$, instead of the amplitude and the phase.

was first discovered, in the context of mechanical oscillators and gravitational-wave detection, by Braginsky (1970), and was first proved with generality by Giffard (1976). However, these amplitude-and-phase limits have long been known in the field of quantum electronics; see, e.g., Serber and Townes (1960).

For a mechanical oscillator of the type to be used in gravitational-wave detection ($m \lesssim 10$ tons, $\omega/2\pi \simeq 1000$ Hz), the standard quantum limit is $\Delta X_j = (\hbar/2m\omega)^{1/2} \gtrsim 1\times10^{-19}$ cm. This is slightly larger than the amplitude changes one expects from a gravitational wave burst due to a supernova explosion in the Virgo cluster of galaxies. Thus amplitude-and-phase measurements of resonant-bar antennas do not look promising for gravitational-wave astronomy [Braginsky (1977); cf. Sec. I of this paper].

"Quantum counting" is an alternative method of measuring a harmonic oscillator. An ideal quantum counter can measure the number operator $\hat{N}$ of the oscillator with complete precision, and can give repeatedly the same result for a sequence of measurements of $\hat{N}$ if no external forces are acting on the oscillator. Equations (2.4) and (2.5) imply

$$\hat{N} = \tfrac{1}{2}(m\omega/\hbar)(\hat{X}_1^2 + \hat{X}_2^2) - \tfrac{1}{2}. \tag{2.17}$$

Hence, a measurement of $\hat{N}$ is equivalent to a measurement of the absolute amplitude $|X| = (X_1^2 + X_2^2)^{1/2}$ of the oscillator. Such a measurement, with complete precision, must leave the phase $\psi = \tan^{-1}(X_2/X_1)$ completely undetermined. In the phase plane the error box for such a measurement is an annulus [Fig. 3(c)]. If one attributes to this error annulus a thickness corresponding to $\delta N = 1$, then its area is $4\pi(\hbar/2m\omega)$—i.e., four times the minimum allowable area.

Quantum counters with high efficiency (high precision) are common devices for photons of optical frequency and higher—e.g., photodiodes and x-ray proportional counters. These counters are all demolition devices; they destroy the photons they count. For photons at infrared frequencies and lower, and for phonons at acoustical frequencies, quantum counters with reasonable efficiency are not yet available. Unruh (1977, 1978) and Braginsky, Vorontsov, and Khalili (1977) have suggested designs of nondemolition devices for measuring the number of photons in a microwave cavity; and Braginsky and Vorontsov (1974) have proposed that one couple such a cavity to a resonant bar, thereby converting bar phonons into cavity photons, measure the number of cavity photons, and thus monitor changes in the number of bar phonons.

Recently the authors (Thorne *et al.*, 1978, 1979) have proposed yet another method of measuring an oscillator: a *"back-action-evading"* measurement of the real part of the complex amplitude, X_1 (or, if one prefers, of the imaginary part X_2). In this method one measures X_1 with high precision; and in the process, in accordance with the uncertainty principle, Eq. (2.9a), one perturbs X_2 by a large amount. In other words, the measuring apparatus is carefully designed so its back action force drives X_2, leaving X_1 largely unscathed; and because X_1 and X_2 are separately conserved, the resulting large uncertainty in X_2 does not feed back onto X_1 as the oscillator evolves. This means that a sequence of high-precision back-action-evading measurements can give the same result for X_1 time and time again.

The error box for a back-action-evading measurement is a long, thin ellipse [Fig. 3(d)], and it becomes a vertical line ($\Delta X_1 = 0$, $\Delta X_2 = \infty$) in the limit of a "perfect measurement." It is instructive to compare the back-action-evading error box [Fig. 3(d)] with the error box for a quick high-precision measurement of x [Fig. 3(a)]. If a first measurement is made at $t = 0$, when $x = X_1$, the subsequent error boxes are qualitatively the same. As the oscillator evolves, these error boxes remain fixed in the (X_1, X_2) coordinate system (X_1 and X_2 are conserved); but they rotate as seen in the $(x, p/m\omega)$ coordinate system. It is this simple fact which makes possible a sequence of arbitrarily accurate measurements of X_1 all giving the same result, and forbids a similar sequence of arbitrarily accurate measurements of x.

In Secs. II.C–II.E we compute, for three types of measurements (amplitude-and-phase, quantum counting, and back-action-evading), the maximum precision with which one can monitor a weak, *classical* force $F(t)$ that drives the oscillator.

C. Monitoring a force by the amplitude-and-phase method

Let an oscillator be driven by a weak classical force $F(t)$, so that its Hamiltonian is

$$\hat{H} = \hat{H}_0 - \hat{x}F(t), \quad \hat{H}_0 = [\text{expression (2.2)}]. \tag{2.18}$$

The classical nature of the force is embodied in the fact that F is a real function of time t rather than an operator. The unitary evolution operator $\hat{U}(t, t_0)$, which governs the evolution of the state vector in the Schrödinger picture, satisfies

$$i\hbar d\hat{U}/dt = \hat{H}(t)\hat{U}, \quad \hat{U}(t_0, t_0) = 1. \tag{2.19}$$

It is straightforward, using the techniques of Sec. 15.9 of Merzbacher (1970), to show that

$$\hat{U}(t, t_0) = \exp[-i(t - t_0)\hat{H}_0/\hbar] \times \exp(-i\beta + \alpha\hat{a}^\dagger - \alpha^*\hat{a}), \tag{2.20a}$$

where

$$\alpha(t, t_0) = (2m\omega\hbar)^{-1/2}\int_{t_0}^{t} F(t')i \times \exp[+i\omega(t' - t_0)]\,dt', \tag{2.20b}$$

$$\beta(t, t_0) = \int_{t_0}^{t} \tfrac{1}{2}i(\alpha^*\dot{\alpha} - \alpha\dot{\alpha}^*)\,dt'. \tag{2.20c}$$

Here a dot denotes a time derivative; $\hat{a}$ and $\hat{a}^\dagger$ are the oscillator's annihilation and creation operators [Eq. (2.3)]; and an asterisk (*) denotes complex conjugation. Notice that α is complex, but β is real. The effect of the force on the oscillator is characterized by the dimensionless quantity α. It will play an important role below.

Now suppose that the oscillator is being studied by a sequence of "amplitude-and-phase" measurements, each of duration $\tau \gtrsim 1/\omega$. How large must the driving force be to produce a measurable change in the oscillator's complex amplitude? Classically the change in

complex amplitude during the time τ is

$$\delta(X_1 + iX_2) = \int_0^\tau \frac{F(t')}{m\omega} ie^{i\omega t'} dt' = \left(\frac{2\hbar}{m\omega}\right)^{1/2} \alpha(\tau, 0) . \quad (2.21)$$

This change is measurable if its absolute magnitude exceeds the diameter of the error box $2(\hbar/2m\omega)^{1/2}$ (standard quantum limit)—i.e., if

$$\text{standard quantum limit: } |\alpha(\tau, 0)| \gtrsim 1 . \quad (2.22a)$$

Note that $|\alpha(\tau, 0)|$ has the physical meaning

$$|\alpha(\tau, 0)|^2 = \frac{1}{\hbar\omega} \cdot \left[\begin{array}{l}\text{total energy that the force } F(t) \text{ would deposit} \\ \text{in a classical oscillator during time } \tau, \text{ if} \\ \text{the oscillator was initially unexcited}\end{array}\right] = \left[\begin{array}{l}\text{the mean number of quanta that the force would} \\ \text{deposit in a quantum-mechanical oscillator, if} \\ \text{the oscillator was initially in its ground state}\end{array}\right] . \quad (2.22b)$$

A fully quantum-mechanical derivation of the measurability criterion (2.22a) proceeds as follows. Assume that a previous, ideal amplitude-and-phase measurement has left the oscillator in a coherent state (Merzbacher, 1970) at time $t=0$:

$$|\psi(0)\rangle = \exp(-\tfrac{1}{2}|\rho|^2 + \rho\hat{a}^\dagger)|0\rangle , \quad (2.23)$$

where ρ is a complex number and $|0\rangle$ is the ground state. This coherent state has

$$\langle \hat{X}_1 + i\hat{X}_2 \rangle = (2\hbar/m\omega)^{1/2}\rho ,$$
$$\Delta X_1 = \Delta X_2 = (\hbar/2m\omega)^{1/2} . \quad (2.24)$$

Then in the Schrödinger picture the oscillator's state at time τ is $|\psi(\tau)\rangle = \hat{U}(\tau, 0)|\psi(0)\rangle$, which by virtue of Eqs. (2.23), (2.20a), and the commutator $[\hat{a}, \hat{a}^\dagger] = 1$ is

$$|\psi(\tau)\rangle = \exp(-i\tau\hat{H}_0/\hbar)\exp[-i\beta + \tfrac{1}{2}(\alpha\rho^* - \alpha^*\rho)] \times \exp[-\tfrac{1}{2}|\alpha + \rho|^2 + (\alpha + \rho)\hat{a}^\dagger]|0\rangle . \quad (2.25)$$

Here $\alpha = \alpha(\tau, 0)$ and $\beta = \beta(\tau, 0)$ are given by Eqs. (2.20b) and (2.20c). This final state, like the initial, is coherent. It has $\langle \hat{X}_1 + i\hat{X}_2 \rangle = (2\hbar/m\omega)^{1/2}(\rho + \alpha)$ and $\Delta X_1 = \Delta X_2 = (\hbar/2m\omega)^{1/2}$. Thus the force $F(t)$ displaces the center of the oscillator's uncertainty circle by

$$\delta\langle \hat{X}_1 + i\hat{X}_2 \rangle = \left(\frac{2\hbar}{m\omega}\right)^{1/2} \alpha(\tau, 0) = \int_0^\tau \frac{F(t')}{m\omega} ie^{i\omega t'} dt' , \quad (2.21')$$

while leaving the size of the error circle unchanged. As expected, this displacement is the same as that derived classically [Eq. (2.21)]; and because the error circle does not change size, the minimum measurable force [Eq. (2.22a)] is also the same. This minimum measurable force has been derived and discussed in the context of gravitation experiments by Braginsky (1970), and with much greater generality by Giffard (1976); see also a recent, very elegant treatment by Hollenhorst (1979).

One might have thought that by a sequence of n measurements one could determine the center of the error circle with accuracy $(\hbar/2m\omega)^{1/2}(1/n)^{1/2}$, and thereby could measure a force $(1/n)^{1/2}$ smaller than Eq. (2.22a). This is not the case because each measurement of precision $(\hbar/2m\omega)^{1/2}$ perturbs the location of the error box by an amount $\gtrsim (\hbar/2m\omega)^{1/2}$. Viewed heuristically, a sequence of n measurements produces a "$\sqrt{n}$" random walk of the error box location that cancels the usual "$1/\sqrt{n}$" improvement of measurement accuracy.

D. Monitoring a force by the quantum-counting method

Next consider quantum-counting measurements of an oscillator on which a classical force is acting. Assume that at time $t=0$ a precise measurement of the number of quanta puts the oscillator into an energy eigenstate $|N\rangle$ with N quanta. Then in the Schrödinger picture the oscillator's state evolves to $|\psi(\tau)\rangle = \hat{U}(\tau, 0)|N\rangle$ during the time interval τ. From Eq. (2.20a), the commutation relation $[\hat{a}, \hat{a}^\dagger] = 1$, and the raising and lowering relations $\hat{a}^\dagger|N\rangle = (N+1)^{1/2}|N+1\rangle$, $\hat{a}|N\rangle = N^{1/2}|N-1\rangle$, one can derive the probability $P(N \to N'; \tau)$ that in the time interval τ the number of quanta changes from N to N':

$$P(N \to N'; \tau) \equiv |\langle N'|\hat{U}(\tau, 0)|N\rangle|^2 = \frac{r!}{s!}[L_r^{(s-r)}(|\alpha|^2)]^2 |\alpha|^{2(s-r)} e^{-|\alpha|^2} , \quad (2.26)$$

where $s \equiv \max(N, N')$, $r \equiv \min(N, N')$, and $L_r^{(n)}(x)$ is the generalized Laguerre polynomial.

The probability that the force has induced any change at all is

$$1 - P(N \to N; \tau) = 1 - e^{-|\alpha|^2}[L_N(|\alpha|^2)]^2 = 1 - e^{-|\alpha|^2}[1 - N|\alpha|^2 + \tfrac{1}{4}N(N-1)|\alpha|^4 - \cdots]^2 . \quad (2.27)$$

This probability is significant if and only if $|\alpha|^2 \gtrsim (N+1)^{-1}$, i.e.,

$$|\alpha(\tau, 0)| \gtrsim (N+1)^{-1/2} . \quad (2.28)$$

This is the criterion for measurability of the force by quantum-counting techniques. It has been derived and discussed by Braginsky (1970) and by Braginsky and Vorontsov (1974); see also the elegant recent treatment by Hollenhorst (1979).

A semiclassical derivation of criterion (2.28) proceeds as follows: Orient the axes of the complex frequency plane so the (unknowable) phase of the initial state is $\psi = 0$; then the initial energy is $E = \frac{1}{2}m\omega^2 X_1^2$; the initial number of quanta is $N = E/\hbar\omega - \frac{1}{2}$

$= \frac{1}{2}[(m\omega/\hbar)X_1^2 - 1]$; the force-induced change in N is $\delta N = (m\omega/\hbar)X_1\delta X_1 = [(2N+1)(m\omega/\hbar)]^{1/2}\delta X_1$, where $\delta(X_1 + iX_2)$ is given by the classical expression (2.21) except for an unknowable phase; to within a factor of order unity, which is fixed by the unknowable phase, $\delta X_1 \simeq (\hbar/m\omega)^{1/2}|\alpha(\tau,0)|$; the criterion of measurability, $\delta N \gtrsim 1$, then comes out to be (2.28), to within a factor of order unity.

Criterion (2.28) implies that, no matter how weak the force F may be, and no matter how short the time interval τ between measurements may be, one can detect the force by preparing the oscillator in a sufficiently energetic initial state (state of sufficiently large N).

When F is large enough to be measured [criterion (2.28) satisfied], then the probability distribution, Eq. (2.26), is not narrowly peaked. Even under the best of circumstances it can reveal

$$\left| \int_0^\tau F(t')e^{i\omega t'}\,dt' \right| = (2m\omega\hbar)^{1/2}|\alpha(\tau,0)|$$

only to within a multiplicative factor of ~3 at the 90% confidence level; cf. Fig. 4. This is far from enough information to permit reconstruction of $F(t)$.

On the other hand, if one had an infinite number of oscillators all coupled to the same classical force (e.g., to a gravitational wave), and all excited to sufficiently high energies, then from the statistics of quantum-counting measurements one could compute the probability distribution (2.26) and from it one could infer $|\alpha(t_2,t_1)|^2$ for any desired t_1 and t_2. Equivalently one could infer $|\int_{t_1}^{t_2} F(t')e^{i\omega t'}\,dt'|^2$—which is sufficient to reveal all details of $F(t)$ except an overall, time-independent sign.

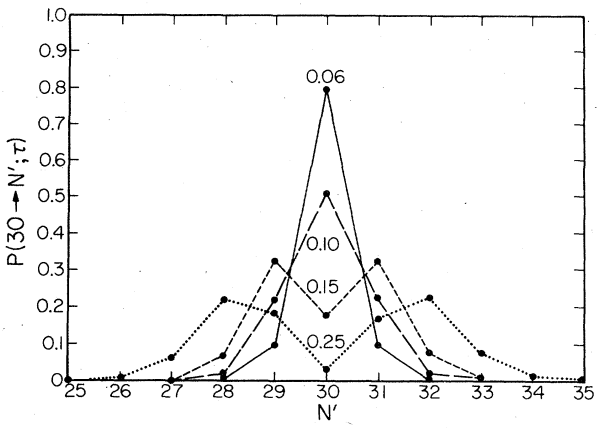

FIG. 4. Harmonic oscillator, initially in an energy eigenstate with $N = 30$ quanta, is driven by a classical force for a time τ. The integrated strength of the force is characterized by the dimensionless number $|\alpha(\tau,0)|$ [Eqs. (2.20b) and (2.22b)]. Here we show the probability $P(30 \to N';\tau)$ that after the force acts the oscillator is in an eigenstate with N' quanta [Eq. (2.26)]. The various probability distributions are labeled by the strength $|\alpha(\tau,0)|$ of the force. Notice that, if one makes a quantum-counting measurement and discovers a transition from $N = 30$ to $N' = 29$, one cannot with confidence determine the strength of the force that acted. One can only conclude that $0.05 \lesssim |\alpha(\tau,0)| \lesssim 0.3$.

E. Monitoring a force by the back-action-evading method

Turn next to our proposed "back-action-evading" method of measuring the $\hat{X}_1$ of an oscillator on which a classical force acts.

In principle, nonrelativistic quantum theory permits $\hat{X}_1$ to be measured "arbitrarily quickly and arbitrarily accurately."[5] By this we mean that, if the oscillator begins the measurement in a near-eigenstate of $\hat{X}_1$, then the measurement can determine the eigenvalue with arbitrary accuracy for any measurement time, no matter how short. We also mean that, regardless of the initial state of the oscillator, the measurement can leave the oscillator in a state arbitrarily close to an eigenstate of $\hat{X}_1$ whose eigenvalue is the measured value ["measurement of the first kind"; Pauli (1958), and Footnote 6].

Such an "arbitrarily quick and accurate" measurement can be achieved by a measuring system which satisfies two requirements: (i) the measuring apparatus must be coupled precisely to $\hat{X}_1$—i.e., it must be coupled to $\hat{X}_1$ and to no other observable of the oscillator; and (ii) the coupling between the measuring apparatus and the oscillator must be arbitrarily strong.[6]

[5]Relativistic quantum theory is not so kind. It places firm constraints on the precision with which certain observables can be measured. For example, the position of a particle (or the X_1 of a harmonic oscillator) cannot be measured with a precision better than the Compton wavelength $\hbar/mc$. Roughly speaking, the reason for this constraint is the following: If one tries to localize a particle within a region smaller than its Compton wavelength, then its momentum uncertainty will be so large that its kinetic energy will be of order its rest mass, and particle-antiparticle pairs will be created. For the X_1 of an oscillator the situation is similar: If one tries to localize X_1 within a Compton wavelength (2×10^{-43} cm if m = 1 ton), then X_2 will be so uncertain that the oscillator's energy will be of order its rest mass. Clearly, this constraint is completely irrelevant for the macroscopic systems considered in this paper.

[6]These two requirements on the measuring apparatus—precise coupling to the measured observable and arbitrarily strong coupling—are also the basic assumptions behind a controversial general "theorem" which asserts, "Nonrelativistic quantum theory permits arbitrarily accurate, *instantaneous* (often called impulsive) measurements of the first kind for any observable." [A measurement of the first kind (Pauli, 1958) is one for which, if the system is in an eigenstate of the measured observable at the instant of the measurement, the result of the measurement is equal to the eigenvalue, with arbitrary accuracy; and regardless of the system's initial state, the measurement leaves it in an eigenstate of the measured observable with the measured eigenvalue.] For a concise review of the literature on this "theorem," see Aharonov and Petersen (1971). This "theorem" is implicit in the viewpoint of Bohr, and it has been championed in recent years by David Bohm. Bohm discusses and gives a proof of the "theorem" in his textbook; see Sec. 22.5 of Bohm (1951). He regards the "theorem" as an immediate consequence of the two requirements on the measuring apparatus. However, one can question the generality of Bohm's proof because of his neglect of the measured system's free Hamiltonian $\hat{H}_0$ during the course of the measurement. In particular, by means of strong forces embodied in the interaction Hamiltonian, the measuring apparatus acts back on variables which do not commute with the measured observable $\hat{A}$. These variables then drive $\hat{A}$ via $\hat{H}_0$, and the resulting disturbance of $\hat{A}$ might preclude (for some observables) arbitrarily accurate measurements even in the [*Footnote continued on next page*].

When requirement (i) is satisfied, $\hat{X}_1$ is completely shielded from noise in the measuring apparatus; then the arbitrarily strong coupling of requirement (ii) can lead to arbitrarily good accuracy for any measurement time, no matter how short. [The crucial property of $\hat{X}_1$—that it is completely shielded from the measuring apparatus when requirement (i) is satisfied—is a general property of "quantum nondemolition observables"; for a precise definition of "quantum nondemolition observable" and a proof of this property, see Sec. IV.]

A skeptic will mistrust this justification of our claim that $\hat{X}_1$ can be measured arbitrarily quickly and accurately. He might worry about the perfection with which one can achieve the time-dependent coupling (Sec. II.F) required for a measurement of $\hat{X}_1$, or he might not believe that $\hat{X}_1$ can be isolated from the measuring apparatus. To alleviate such worries, we describe in Sec. III.B a *Gedanken* experiment which shows that arbitrarily quick and accurate measurements can be made.

Of course *in practice* there are limits to the quickness and precision with which $\hat{X}_1$ can be measured—limits imposed by the strengths of real materials, voltage breakdown in capacitors, etc. In Paper II we discuss some of these practical issues. Until then, however, we restrict attention to limits of principle which are imposed by nonrelativistic quantum mechanics. In this context the crucial point is that, whereas the uncertainty principle of nonrelativistic quantum theory places severe restrictions on the accuracy of amplitude-and-phase measurements, it places no restriction whatsoever on the speed or accuracy of measurements of $\hat{X}_1$.

We now compute the precision with which one can monitor a classical force $F(t)$ by back-action-evading measurements of $\hat{X}_1$. Our computation is carried out in the Heisenberg picture. Suppose that an initial precise measurement of $\hat{X}_1$, at time $t=t_0$, gives a value ξ_0 and leaves the oscillator in the corresponding eigenstate $|\xi_0\rangle$ of $\hat{X}_1$. (The spectrum of $\hat{X}_1$, like the spectra of $\hat{x}$ and $\hat{p}$, is continuous; thus ξ_0 can be any real number.) As time passes the state of the oscillator remains fixed in the Heisenberg picture, $|\psi(t)\rangle\rangle \equiv |\xi_0\rangle$, but $\hat{X}_1$ evolves:

$$\frac{d\hat{X}_1}{dt} = -\frac{i}{\hbar}[\hat{X}_1, \hat{H}] + \frac{\partial \hat{X}_1}{\partial t} = -\frac{F(t)}{m\omega}\sin\omega t \tag{2.29}$$

[Eqs. (2.18), (2.2), (2.5a), and (2.1)]. Integrating this equation, we obtain

$$\hat{X}_1(t) = \hat{X}_1(t_0) - \int_{t_0}^{t} \frac{F(t')}{m\omega}\sin(\omega t')\,dt' . \tag{2.30}$$

Because $|\psi(t)\rangle\rangle = |\xi_0\rangle$ is an eigenstate of $\hat{X}_1(t_0)$, and because $\int_{t_0}^{t}[F(t')/m\omega]\sin(\omega t')\,dt'$ is a real number, $|\psi(t)\rangle\rangle$ is also an eigenstate of $\hat{X}_1(t)$ with eigenvalue

$$\xi(t, t_0) = \xi_0 - \int_{t_0}^{t} \frac{F(t')}{m\omega}\sin(\omega t')\,dt' . \tag{2.31}$$

A precise measurement of $\hat{X}_1(t)$ at time t must then yield this eigenvalue $\xi(t, t_0)$ and must leave the state of the oscillator unchanged (except for overall phase).

This remarkable fact—that even when a classical force is acting, successive perfect measurements of $\hat{X}_1$ leave the oscillator's state unchanged—means that perfect measurements of $\hat{X}_1$ are "quantum nondemolition" in a stronger sense than quantum-counting measurements can ever be. In the quantum-counting case the classical force drives the oscillator away from eigenstates of the measured operator $\hat{N}$, and a subsequent perfect measurement then "demolishes" the oscillator's evolved state—i.e., it "reduces the wave function" back into an eigenstate of $\hat{N}$. Perfect quantum-counting experiments are truly nondemolition only in the absence of an external driving force.

By a sequence of arbitrarily quick and accurate back-action-evading measurements of $\hat{X}_1$ one can monitor, in principle, the precise time evolution of the oscillator's eigenvalue $\xi(t, t_0)$ [Eq. (2.31)]; and from $\xi(t, t_0)$ one can compute the precise time evolution of the driving force (signal):

$$F(t) = -(m\omega\, d\xi/dt)/(\sin\omega t) . \tag{2.32}$$

Of course, in the realistic case of imperfect measurements, the inferred $F(t)$ will be highly inaccurate at times $t \simeq n\pi/\omega$, when $\sin\omega t \simeq 0$. However, when the force is produced by a classical *field* (e.g., a gravitational or electromagnetic wave) whose wavelength is long compared to the size of the measuring apparatus, one can couple two different oscillators to F. On the first oscillator one can measure $\hat{X}_1$ getting $\xi(t, t_0)$, and on the second one can measure $\hat{X}_2$ getting

$$\zeta(t, t_0) \equiv \zeta_0 + \int_{t_0}^{t} \frac{F(t')}{m\omega}\cos(\omega t')\,dt' .$$

One can infer $F(t)$ independently from the two measurements, and the accuracies of the two methods will be complementary: The second is good at $t = n\pi/\omega$ when the first is bad; the first is good at $t = (n+\frac{1}{2})\pi/\omega$ when the second is bad.

This technique of measuring $\hat{X}_1$ on one oscillator and $\hat{X}_2$ on another completely circumvents the uncertainty principle. In the complex amplitude plane the vertical error line associated with $\hat{X}_1$ (first oscillator), and the

---

limit of zero measurement time. To prove the "theorem" in a particular case, one must include the effects of $\hat{H}_0$ and one must show that the measurement error goes to zero in an appropriate limit where the coupling strength goes to infinity and the measurement time goes to zero. In general, the error can be made to go to zero *only* in the limit of an instantaneous measurement. Fortunately, for the observables considered in this paper (such as the position of a free particle or harmonic oscillator) the theorem is undoubtedly true. Indeed, for "quantum nondemolition observables" the theorem holds in the stronger form given in the text for $\hat{X}_1$ (arbitrarily accurate measurements even for *nonzero* measurement times). The "theorem" has long been controversial because it implies (in its stronger form) that the energy of a system can be measured arbitrarily quickly and accurately, in violation of a common misinterpretation of the energy-time uncertainty relation. [For a specific *Gedanken* experiment that proves the possibility of arbitrarily quick and accurate energy measurements, see Aharonov and Bohm (1961, 1964). The latter is a valid special case of Bohm's (1951) proof of the general "theorem."] The misinterpretation of $\Delta E \Delta t \gtrsim \hbar$ has generated so much confusion in the physics community that even Von Neumann (1932; Sec. V.1) regarded it as a counterexample to the "theorem."

horizontal error line associated with $\hat{X}_2$ (second oscillator), intersect in a point. This point moves, under the action of $F(t)$, in precisely the same manner as the system point of a single *classical* oscillator driven by $F(t)$; see Fig. 5.

That measurements of $\hat{X}_1$ can reveal all details of $F(t)$, while quantum-counting measurements cannot, is intimately connected with the fact that measurements of $\hat{X}_1$ are quantum nondemolition even in the presence of a classical force while quantum-counting measurements are not. For further discussion see Sec. IV.

Of course, in practice there are limits to the accuracy with which back-action evasion can monitor an external force $F(t)$. The most serious limits arise from Nyquist noise in the oscillator, and from constraints on the strength of coupling of real transducers to the oscillator—constraints due to the finite strengths of real materials, voltage breakdown in real capacitors, and superconducting breakdown in real circuits; see Paper II. Less serious in practice, but important in principle, are limits due to special relativistic effects (see footnote 5), and a limit due to the quantum-mechanical properties of any real external force.

The latter limit, which we shall call the *"real quantum limit,"* arises when one is monitoring the external force F so accurately that one discovers it is not classical, but rather is being produced by a boson system with a finite occupation number per quantum-mechanical state. The magnitude of this real quantum limit on the force F is a function of the strength of coupling of the boson system to the oscillator: The weaker the coupling, the smaller will be the magnitude of the force at which the system's quantum properties are felt. To quantify the strength of coupling unambiguously, consider the case where the boson system produces a force whose duration is approximately one cycle of the oscillator—i.e., a broadband force with bandwidth $\Delta\omega \sim \omega$. Then consider all quantum states associated with this driving force (e.g., if the force is produced by electromagnetic or gravitational radiation, consider all states in the beam pattern of the antenna with frequencies in the range $\Delta\omega \sim \omega$). Let $\bar{n}_{\rm SQL}$ be the average occupation number of these states when the force is just strong enough to be detectable in one cycle by amplitude-and-phase methods [force at level of "standard quantum limit," Eqs. (2.16) and (2.22a)]. Then $\bar{n}_{\rm SQL}$ characterizes the strength of coupling of the oscillator to the boson system. In the special case of an antenna for electromagnetic or gravitational waves, one can show that

$$\bar{n}_{\rm SQL} \simeq \lambda\!\!\!^{-2}/\sigma_0 \simeq 10^{38} \text{ for resonant-bar gravitational-wave antennas} \qquad (2.33)$$

(cf. Sec. I.B). Here $\lambda\!\!\!^{-}=c/\omega$ is the reduced wavelength of the waves and σ_0 is the cross section of the antenna [equal to $\omega^{-1}\int\sigma(\omega')d\omega'$ where the integral is over the antenna's resonance and $\sigma(\omega')$ is the cross section at frequency ω'; cf. Chap. 37 of Misner, Thorne, and Wheeler (1973)].

Now consider an arbitrary force and a measurement of X_1 that lasts a time $\tau \gtrsim \omega^{-1}$. The real quantum limit for such a measurement is reached at a level that is smaller than the standard quantum limit by $(\omega\tau/\bar{n}_{\rm SQL})^{1/2}$:

$$\left.\begin{array}{l}\text{real}\\ \text{quantum}\\ \text{limit}\end{array}\right\} \begin{array}{l}\Delta X_1 \simeq (\hbar/2m\omega)^{1/2}(\omega\tau/\bar{n}_{\rm SQL})^{1/2},\\ |\alpha(\tau,0)| \simeq (\omega\tau/\bar{n}_{\rm SQL})^{1/2}.\end{array} \qquad (2.34)$$

If one were monitoring the force $F(t)$ by back-action-evading techniques at this level of accuracy, one's measurements would be sensitive to zero-point (vacuum) fluctuations in the system that produces the force F.

Henceforth, as previously, we shall ignore these issues and shall regard the force $F(t)$ as truly classical ($\bar{n}_{\rm SQL}=\infty$).

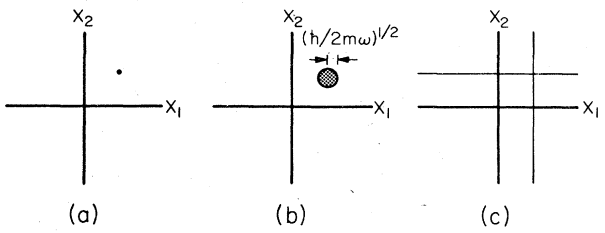

FIG. 5. (a) Classical harmonic oscillator is described by a single "system point," which moves about in the complex amplitude plane in response to an external driving force [Eq. (2.21)]. (b) Quantum-mechanical oscillator in a coherent state is described by a minimum-uncertainty wave packet. In the absence of measurements the center of that wave packet moves about in the complex amplitude plane, in response to an external driving force, with precisely the same motion as the system point of the classical oscillator [Eq. (2.21′)]. However, it is impossible to measure that motion more precisely than $\Delta X_1 = \Delta X_2 = (\hbar/2m\omega)^{1/2}$. (c) Two quantum-mechanical oscillators, one in an eigenstate of $\hat{X}_1$, the other in an eigenstate of $\hat{X}_2$, are described by two orthogonal error lines in the complex amplitude plane. Under the action of an external driving force the intersection of the two error lines moves in exactly the same manner as the system point of a classical oscillator. In principle, this motion can be measured with complete precision, and without perturbing the error lines, by means of back-action-evading measurements.

F. Interaction Hamiltonians for back-action-evading measurements of $\hat{X}_1$

1. Continuous two-transducer measurements

A back-action-evading measurement of $\hat{X}_1$ is made by (i) coupling the oscillator to a measuring apparatus which produces an output large enough to be essentially classical, (ii) reading out the output of the measuring apparatus, and (iii) inferring a value for X_1 from that output. The coupling of the oscillator to the measuring apparatus is embodied, mathematically, in the "interaction part" of the Hamiltonian $\hat{H}_I$. To prevent back action of the measuring apparatus on $\hat{X}_1$ it is necessary that $\hat{H}_I$ commute with $\hat{X}_1$. To make the measurement of very small X_1's experimentally feasible, it is advantageous to use a linear coupling of the measuring apparatus to the oscillator's position and momentum. These constraints of linear coupling and commutation with $\hat{X}_1$ force $\hat{H}_I$ to have the form

$$\hat{H}_I = K\hat{X}_1\hat{Q} = K[\hat{x}\cos\omega t - (\hat{p}/m\omega)\sin\omega t]\hat{Q}. \qquad (2.35)$$

Here K is a "coupling constant" that may be time dependent, and $\hat{Q}$ is an operator (observable) of the measuring apparatus. ($\hat{Q}$ commutes with all the oscillator observables.) The total Hamiltonian for the coupled system consisting of the oscillator, the measuring apparatus, and the classical driving force has the form

$$\hat{H} = \hat{H}_0[\text{Eq. (2.2)}] - F(t)\hat{x} + \hat{H}_I[\text{Eq. (2.35)}] + \hat{H}_M . \qquad (2.36)$$

Here $\hat{H}_M$ is the Hamiltonian of the measuring apparatus—i.e., it is the part of the Hamiltonian that depends only on measuring apparatus observables.

When K is time independent, the interaction Hamiltonian (2.35) can be realized as follows: One couples the oscillator to a coordinate (x) transducer, and one sinusoidally modulates the transducer output at the frequency ω of the oscillator; one also couples the oscillator to a momentum (p) transducer and modulates its output sinusoidally with a phase which leads that of the coordinate transducer by a quarter cycle; one adds the two outputs and sends the sum into an amplifier. (The sinusoidal modulations must be produced by a classical signal generator—e.g., another oscillator with the same frequency as the primary oscillator, vibrating in a large-amplitude coherent state.) Specific designs for this type of measuring apparatus will be described in Appendix B, and in Paper II. In Sec. III.B we shall see that, if the coupling constant K is made arbitrarily large, then in principle the measurement of $\hat{X}_1$ can be made arbitrarily quickly and arbitrarily accurately.

We shall characterize such measurements as "continuous two-transducer measurements."

2. Stroboscopic measurements

If one is willing to make measurements only twice per cycle, then one can avoid the necessity for both coordinate and momentum transducers. In particular, if one pulses on the coupling at times $t = n\pi/\omega$, so $K = K_0\delta(\sin\omega t)$, then the interaction Hamiltonian (2.35) becomes

$$\hat{H}_I = K_0 \cos\omega t\, \delta(\sin\omega t)\, \hat{x}\hat{Q} = \frac{K_0}{\omega}\hat{x}\hat{Q}\sum_n (-1)^n \delta\left(t - \frac{n\pi}{\omega}\right), \qquad (2.37a)$$

which requires only a coordinate transducer for its realization. [The factor $(-1)^n$, i.e., the sign change in the coupling between even and odd pulses, compensates for the sign change in the relation between $\hat{x}$ and $\hat{X}_1$: $\hat{X}_1 = (-1)^n\hat{x}$.] If one pulses on the coupling at $t = (n+\frac{1}{2})\pi/\omega$, then

$$\hat{H}_I = (K_0/m\omega)[-\sin\omega t\, \delta(\cos\omega t)]\hat{p}\hat{Q}, \qquad (2.37b)$$

which requires only a momentum transducer. The possibility of such pulsed measurements was discovered independently and simultaneously by Zimmermann in our research group (see Thorne *et al.*, 1978), and by Braginsky, Vorontsov, and Khalili (1978) in Moscow. Braginsky *et al.* call such measurements "stroboscopic."

Stroboscopic measurements with the interaction Hamiltonian (2.37a) can be described semiclassically as follows: One measures the oscillator's coordinate $x = X_1$ at $t = 0$, obtaining a precise value ξ_0 and in the process giving the momentum a huge unknowable uncertainty-principle kick. The kick causes x to evolve in an unknown way. However, because the oscillator's period is independent of its amplitude, after precisely a half-cycle x must be precisely equal to $-\xi_0$ in the absence of an external force, or equal to

$$-\xi\left(\frac{\pi}{\omega}, 0\right) = -\left(\xi_0 - \int_0^{\pi/\omega} \frac{F(t')}{m\omega}\sin\omega t'\, dt'\right) \qquad (2.38)$$

in the presence of a classical force F [cf. Eq. (2.31)]. At $t = \pi/\omega$ a second pulsed measurement is made, giving precisely this value for $x = -X_1$, and again kicking the momentum by a huge, unknowable amount. Subsequent pulsed measurements at $t = n\pi/\omega$ give values

$$x = (-1)^n\xi\left(\frac{n\pi}{\omega}, 0\right) = (-1)^n\left(\xi_0 - \int_0^{n\pi/\omega} \frac{F(t')}{m\omega}\sin\omega t'\, dt'\right), \qquad (2.39)$$

which are unaffected by the unknown kick of each measurement.

In the Schrödinger picture of quantum mechanics these stroboscopic measurements are described as follows: A precise measurement of $\hat{x}$ at $t = 0$ gives the value ξ_0 and collapses the oscillator's wave function $\psi(x, 0)$ into an arbitrarily narrow function peaked at ξ_0—i.e., $\psi(x, 0) \simeq [\delta(x - \xi_0)]^{1/2}$. Immediately after the measurement the wave function $\psi(x, t)$ spreads out over all space; but as t approaches π/ω, ψ gathers itself into an arbitrarily narrow function again, now centered on $x = -\xi(\pi/\omega, 0)$ [Eq. (2.38)]. A precise measurement of $\hat{x}$ at this time gives this precise value and leaves the oscillator's state unchanged except for phase (no collapse of wave function; quantum nondemolition measurement). Just before each subsequent measurement ($t = n\pi/\omega$) the wave function again collects itself into an arbitrarily narrow function, and a perfect nondemolition measurement can again be made.

In practice, of course, no measurement can be made perfectly. The following simple argument reveals the limit of accuracy for stroboscopic measurements which require a finite time $2\Delta t$, or which are made at times that differ by Δt from precise half-cycle timing. (A more rigorous calculation gives the same limit.) Let σ be the precision of such a measurement at $t \simeq 0$. Then immediately after the measurement the oscillator's wave function must have variances $\Delta x_0 = \sigma$, $\Delta p_0 \geq \hbar/2\sigma$. The next measurement will have optimal accuracy only if the first measurement has put the wave function into a minimum-uncertainty wave packet ($\Delta p_0 = \hbar/2\sigma$). Then, as time passes, the variances of $\hat{x}$ and $\hat{p}$ feed each other so that, at the time $t = (\pi/\omega \pm \Delta t)$ of the next measurement,

$$\Delta x = [(\Delta x_0)^2\cos^2\omega t + (\Delta p_0/m\omega)^2\sin^2\omega t]^{1/2} = [\sigma^2 + (\hbar\Delta t/2m\sigma)^2]^{1/2} \quad \text{for } \Delta t \ll \pi/\omega . \qquad (2.40)$$

This is the minimum possible uncertainty for the next measurement. It is minimized (optimal strategy!) by setting $\sigma = (\hbar\Delta t/2m)^{1/2}$, which gives

$$\Delta X_1 = \Delta x = [(\hbar/m\omega)(\omega\Delta t)]^{1/2} \qquad (2.41)$$

for the best possible accuracy of stroboscopic measurements with timing imperfections Δt. This result

has been derived independently by Thorne *et al.* (1978) and by Braginsky, Vorontsov, and Khalili (1978).

3. Continuous single-transducer measurements

Return now to continuous measurements. The more rapidly one seeks to measure, the larger must be the coupling constant K. This fact will be quantified in Eq. (3.21) below and in Paper II. In many situations, practical considerations will force K to be so small that measurements of the desired accuracy will require a time τ far longer than one cycle. In such cases, as in stroboscopic measurements, one can avoid the use of two transducers. For example, one can construct an interaction Hamiltonian of the form

$$\hat{H}_I = K\hat{Q}\hat{x}\cos\omega t = \tfrac{1}{2}K\hat{Q}(\hat{X}_1 + \hat{X}_1\cos 2\omega t + \hat{X}_2\sin 2\omega t) \tag{2.42a}$$

or

$$\hat{H}_I = (K/m\omega)\hat{Q}\hat{p}\sin\omega t = -\tfrac{1}{2}K\hat{Q}(\hat{X}_1 - \hat{X}_1\cos 2\omega t - \hat{X}_2\sin 2\omega t)\,; \tag{2.42b}$$

cf. Eq. (2.6). The first of these is achieved by a coordinate transducer with sinusoidally modulated output; the second, by a momentum transducer with modulated output. Measurements with such Hamiltonians we shall call "continuous single-transducer measurements."

In such single-transducer measurements, the apparatus which follows the transducer must average over a time $\bar{\tau} \gg 2\pi/\omega$ in producing its output—i.e., it must contain a "low-pass filter" with high-frequency cutoff at $\omega_{\max} \simeq \pi/\bar{\tau} \ll \omega$. Then the sinusoidal output due to the sinusoidal terms in $\hat{H}_I$ [Eqs. (2.42)] will average away to near zero. To free $\hat{X}_1$ from back-action forces of the measuring apparatus, one must ensure that the back-action forces have negligible Fourier components at frequency 2ω. This can be arranged, for example, by placing a low-pass filter between the transducer and the subsequent apparatus. See Paper II for full details.

Such a continuous single-transducer back-action-evading measurement is similar to a lock-in amplifier. In the lock-in amplifier a slowly changing signal S is given an initial "carrier" modulation $S\cos\omega t$ before it acquires (through amplification and other signal processing) a noise N. The noisy signal $S\cos\omega t + N$ is then subjected to "phase-sensitive detection"—i.e., it is multiplied by $\cos\omega t$ and then is sent through a low-pass filter to give a signal $\frac{1}{2}S$ which is nearly free of the noise N. By contrast, in our back-action-evading measurement of an oscillator, the oscillator itself provides the initial modulation of the "signal" X_1 to produce a "carrier" $x = X_1\cos\omega t + X_2\sin\omega t$—which then enters the signal-processing apparatus through a transducer. The subsequent modulation and filtering of the carrier are identical to the phase-sensitive detection of the lock-in amplifier, except for this key difference: In the lock-in amplifier the phase-sensitive detection follows amplification, and its purpose is to remove from the signal the noise inserted during signal processing; in our back-action-evading measurement the phase-sensitive detection precedes amplification, and its purpose is to make one's measurement insensitive to the noisy back action of the amplifier on the oscillator, which was the source of the initial modulation. (For comments on the related issue of the similarity between our back-action-evading measurements and the operation of a degenerate parametric amplifier, see footnote 2.)

The modulation in our single-transducer interaction Hamiltonian (2.42) need not be sinusoidal, nor need it be at the oscillator frequency. A variety of other types of modulation will do the job—if they are accompanied by appropriate filters placed between the transducer and the subsequent apparatus. For details see Paper II; for practical examples see Thorne *et al.* (1979) and Braginsky *et al.* (1980).

It seems likely to us that (at least for gravitational-wave detection) the most practical type of back-action-evading measurement will be continuous single-transducer measurements. A practical variant of a single transducer measurement involves only modest modifications of standard "amplitude-and-phase" electronic techniques. The essential feature of a practical design is the following: The modulation of the transducer output and the subsequent filtering must precede amplification. It is this that allows $\hat{X}_1$ to evade the amplifier's back-action noise.

In Appendix D we show that continuous single-transducer back-action-evading measurements with averaging times $\bar{\tau} \gg 2\pi/\omega$ are capable of accuracies

$$\Delta X_1 \simeq (\hbar/2m\omega)^{1/2}(1/\omega\bar{\tau})^{1/2} \tag{2.43}$$

[Eqs. (D16) and (D21)]. (Appendix D is best read after Sec. III and Appendix C.) Paper II will discuss practical limitations on modulated measurements—including limitations due to finite strength of coupling K; see also Eqs. (D19)–(D21) of this paper.

Continuous single-transducer back-action-evading measurements of $\hat{X}_1$ are analogous to the single-transducer quantum-counting measurements proposed by Unruh (1977, 1978) and by Braginsky, Vorontsov, and Khalili (1977). The Unruh-Braginsky interaction Hamiltonian has the form

$$\hat{H}_I = K\hat{Q}\hat{x}^2 = \tfrac{1}{2}K\hat{Q}[(2\hbar/m\omega)(\hat{N} + \tfrac{1}{2}) + (\hat{X}_1^2 - \hat{X}_2^2)\cos 2\omega t + (\hat{X}_1\hat{X}_2 + \hat{X}_2\hat{X}_1)\sin 2\omega t] \tag{2.44}$$

[cf. Eqs. (2.17) and (2.5)], which is analogous to our equation (2.42); and they measure $\hat{N}$ by averaging over a time $\tau \gg 2\pi/\omega$.

G. Zero-frequency limit of back-action-evading measurements

In the limit $\omega \to 0$ a mechanical oscillator becomes a "free mass," and the real and imaginary parts of the complex amplitude become

$$\hat{X}_1 = \hat{x} - (\hat{p}/m)t\,, \quad m\omega\hat{X}_2 = \hat{p}\,. \tag{2.45}$$

For a free mass these quantities, $\hat{p}$ and $\hat{x} - (\hat{p}/m)t$, are conserved in the absence of external forces; and one can monitor a classical external force by "back-action-evading measurements" of either of these quantities.

To measure $\hat{p}$ requires only a momentum trans-

ducer—i.e., a transducer whose interaction Hamiltonian is

$$\hat{H}_I = K\hat{p}\hat{Q} . \tag{2.46}$$

To measure $\hat{X}_1$ requires both a position transducer and momentum transducer

$$\hat{H}_I = K\hat{X}_1\hat{Q} = K\hat{x}\hat{Q} - K\hat{p}\hat{Q}t/m . \tag{2.47}$$

As in the case of a harmonic oscillator, so also for a free mass, a measurement of $\hat{X}_1$ or $\hat{p} = m\omega\hat{X}_2$ can be arbitrarily quick and arbitrarily accurate in principle (as long as one ignores issues of strengths of materials, relativistic effects, etc.). This we demonstrate in Sec. III.

III. *GEDANKEN* EXPERIMENTS FOR ARBITRARILY QUICK AND ACCURATE BACK-ACTION-EVADING MEASUREMENTS OF $\hat{X}_1$ OR $\hat{X}_2$

Here we describe and analyze *Gedanken* experiments by which, in principle, one can measure arbitrarily quickly and accurately (i) the momentum $\hat{p} = \lim_{\omega\to 0}(m\omega\hat{X}_2)$ of a free mass, and (ii) the real part $\hat{X}_1$ of the complex amplitude of a harmonic oscillator. Throughout this section, as above, the phrase "in principle one can measure arbitrarily quickly and accurately" implicitly contains the caveat "within the framework of nonrelativistic quantum mechanics and ignoring constraints due to strengths of materials, voltage breakdown in capacitors, relativistic effects, etc." Consequently, in this section and related appendixes we shall, without further comment or shame, take limits in which sizes of capacitors go to infinity, energies in electromagnetic frequency generators (clocks) go to infinity, etc. To alleviate queasiness caused by this cavalier approach, we shall administer a strong dose of practical constraints in Paper II.

In this section we shall first (subsection A) discuss measurements of free masses, and then (subsection B) measurements of oscillators.

A. Measurements of a free mass

1. Standard quantum limit

Gedanken experiments described in the literature suggest a possible limit

$$\text{standard quantum limit: } (\Delta F)_{\min} \simeq (m\hbar/\tau^3)^{1/2} \tag{3.1}$$

on the accuracy with which one can measure a weak classical force F acting on a free mass m, with a measurement of duration τ.

This "standard quantum limit" is correct and unavoidable (Braginsky and Vorontsov, 1974) if one tries to study F by measurements of the mass's position [analog of "amplitude-and-phase" method for an oscillator; cf. Eqs. (3.1) and (2.22)]. An initial position measurement of precision Δx_i produces, by the position-momentum uncertainty relation, a variance $\Delta p \geq \Delta p_{\min} = \hbar/2\Delta x_i$ in the mass's initial momentum, which in turn produces the following variance of position after a time τ:

$$\Delta x(\tau) \geq \left[(\Delta x_i)^2 + \left(\frac{\Delta p_{\min}}{m}\right)^2 \tau^2\right]^{1/2} = \left[(\Delta x_i)^2 + \left(\frac{\hbar\tau}{2m\Delta x_i}\right)^2\right]^{1/2} \geq \left(\frac{\hbar\tau}{m}\right)^{1/2} \tag{3.2}$$

("standard quantum limit for free-mass position"). In this same time τ a constant force F produces a change of position $\delta x = \frac{1}{2}(F/m)\tau^2$. Comparing the signal δx with the noise (3.2), we obtain the standard quantum limit (3.1) on the force F, to within a factor 2. A laser-interferometer detector for gravitational waves is an example of a system which studies weak classical forces by position measurements, and which is therefore subject to the constraint (3.1); see, e.g., Drever *et al.* (1977) or Edelstein *et al.* (1978). For laser detectors this constraint is a serious potential problem at low gravitational-wave frequencies, $f \lesssim 1$ Hz.

Another measuring system that is subject to the constraint (3.1) is a "velocity sensor." By "velocity sensor" we mean a measuring system in which, viewed classically, the velocity $\dot{x}$ of the mass m produces an emf in a circuit, and the effects of that emf are measured using a voltage or current or charge amplifier. An idealized simple-minded version of such a sensor is shown in Fig. 6(a). For that sensor or any "velocity sensor," the Lagrangian of the entire system, with amplifier disconnected, has the form

$$\mathcal{L} = \tfrac{1}{2}m\dot{x}^2 + \tfrac{1}{2}L\dot{Q}^2 - (1/2C)Q^2 - Km\dot{x}Q + Fx . \tag{3.3}$$

Here F is the force on m, which one seeks to measure; Q is the charge that has flowed onto the upper plate of the capacitance C; $\dot{Q}$ is the current in the circuit; and for the system of Fig. 6(a) the coupling constant is $K = aB/mc$, where a is the height of the magnetic-field

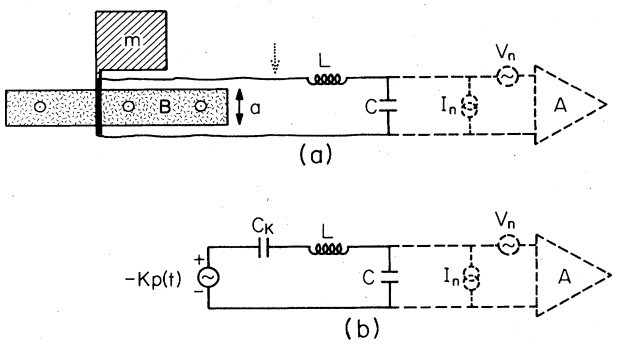

FIG. 6. (a) Idealized velocity sensor. The "free mass" m has a wire (dark vertical bar) rigidly attached to it. The wire is hooked up to an LC circuit; and it passes through a region of uniform magnetic field (stippled region). The velocity $\dot{x}$ of the mass produces an emf $Ba\dot{x}/c$ in the LC circuit. During a measurement one either attaches a voltage amplifier in parallel with the capacitance C (dashed part of figure) and makes C as small as possible (open circuit), or one attaches a charge or current amplifier in series (not shown) and makes C as large as possible (short circuit). In any case, to achieve minimum noise one makes the stray inductance L as small as possible. As discussed in Sec. III.A.2, one can turn this velocity sensor into a momentum sensor by inserting a negative capacitance $-C_K = -1/mK^2 = -mc^2/(aB)^2$ at the location indicated by a dotted arrow. (b) Equivalent circuit for the velocity sensor of (a); see text for discussion.

region, B is the field strength, and c is the speed of light. The generalized momenta of this system are

$$p=\frac{\partial \mathcal{L}}{\partial \dot{x}}=m(\dot{x}-KQ)\,,\quad \Pi=\frac{\partial \mathcal{L}}{\partial \dot{Q}}=L\dot{Q}\,; \tag{3.4}$$

and the Hamiltonian $H=p\dot{x}+\Pi\dot{Q}-\mathcal{L}$, after quantization, is

$$\hat{H}=\frac{\hat{p}^2}{2m}+\frac{\hat{\Pi}^2}{2L}+\tfrac{1}{2}\left(mK^2+\frac{1}{C}\right)\hat{Q}^2+K\hat{p}\hat{Q}-F\hat{x}\,. \tag{3.5}$$

Note that *the velocity coupling $-Km\dot{x}Q$ in the Lagrangian is equivalent to a momentum coupling $K\hat{p}\hat{Q}$ in the Hamiltonian plus a capacitance $C_K\equiv 1/mK^2$ in the readout circuit.* It is the capacitance C_K which prevents such a velocity sensor even in principle from monitoring the momentum p and force F with arbitrary speed and accuracy.

A semiclassical derivation of the quantum limit (3.1) for such a velocity sensor proceeds as follows: If the mass is initially in an eigenstate (or near eigenstate) of $\hat{p}$ with eigenvalue p_0, then the form (3.5) of the Hamiltonian guarantees it will remain in an eigenstate of $\hat{p}$ but with eigenvalue $p(t)=p_0+Ft$. (Here F is assumed constant, for simplicity.) Figure 6(b) is then an equivalent circuit for the measuring apparatus. Simple analysis of this circuit, with voltage amplifier included, shows that the output $\tilde{V}_a$ of the amplifier at frequency f is

$$\tilde{V}_a=A(f)\left(\frac{-K\tilde{p}+\tilde{I}_n(-i2\pi fL+1/-i2\pi fC_K)}{1+C/C_K-(2\pi f)^2CL}+\tilde{V}_n\right). \tag{3.6a}$$

Here a tilde denotes a Fourier transform, and for simplicity we have assumed that the amplifier has infinite input impedance and that initially there is zero charge on the capacitor and zero current through the inductor. For a quick measurement of duration τ (frequency and bandwidth $f\sim\Delta f\sim 1/2\tau$), the signal-to-noise ratio (SNR) is optimized by setting $C=L=0$; then

$$\mathrm{SNR}\sim\frac{K(p_0+\frac{1}{2}F\tau)}{\{[S_V(f)+S_I(f)/(2\pi fC_K)^2]\Delta f\}^{1/2}}\,. \tag{3.6b}$$

Here $S_V(f)$ is the spectral density of the amplifier's voltage noise V_n and $S_I(f)$ is the spectral density of its current noise I_n. The Heisenberg uncertainty principle constrains the noise temperature of the amplifier to be $T_n\geq 2\pi\hbar f/k\ln 2$ (Weber, 1959; Heffner, 1962) which, by virtue of Eq. (12.33) of Robinson (1974), is equivalent to the constraint

$$S_V(f)S_I(f)\gtrsim(4\pi\hbar f)^2 \tag{3.7}$$

[cf. Eq. (2.11)]. The ratio S_V/S_I can be adjusted by preceding the amplifier with a transformer. The optimal SNR occurs when $S_V/S_I=1/(2\pi fC_K)^2$, which—together with $f\sim\Delta f\sim 1/2\tau$—gives

$$(\mathrm{SNR})_{\mathrm{opt}}\sim(C_K\tau/\hbar)^{1/2}K(p_0+\tfrac{1}{2}F\tau)\,. \tag{3.8}$$

Since $C_K=1/mK^2$, this optimal SNR does not improve as $K\to\infty$. In fact, independent of K the minimum detectable force (SNR$\simeq$ 1) is the "standard quantum limit" (3.1). For the case of a charge or current amplifier in series with the circuit (and for optimization of the circuit impedances to $C=\infty$, $L=0$), a similar analysis gives the same limit.

Wagoner, Will, and Paik (1979) have proposed a design for a free-mass gravitational-wave detector which makes use of a velocity sensor. Their technique for coupling the circuit to the mass is essentially equivalent to the technique shown in Fig. 6(a), but is a more practical variant of it. Their Lagrangian has the standard velocity-sensor form of Eq. (3.3), and therefore its performance can *never* exceed the "standard quantum limit" (3.1).

2. Momentum sensors can be arbitrarily quick and accurate

From a velocity sensor such as that in Fig. 6 one can construct a momentum sensor by inserting into the circuit a capacitor with negative capacitance $-C_K=-1/mK^2$. A negative capacitor is not a common electronic component. Nevertheless, such capacitors can exist in principle, and in principle their internal noise can be made negligible; see Appendix A for details.

The momentum sensing system, which one obtains from the velocity sensor of Eq. (3.3) by inserting the negative capacitance $-C_K=-1/mK^2$, has the Lagrangian

$$\mathcal{L}=\tfrac{1}{2}m\dot{x}^2+\tfrac{1}{2}L\dot{Q}^2+(\tfrac{1}{2}mK^2-1/2C)Q^2-Km\dot{x}Q+Fx\,. \tag{3.9}$$

Its velocities and momenta are related by Eq. (3.4), and its quantized Hamiltonian is Eq. (3.5) with negative capacitance inserted:

$$\hat{H}=\hat{p}^2/2m+\hat{\Pi}^2/2L+(1/2C)\hat{Q}^2+K\hat{p}\hat{Q}-F\hat{x}\,. \tag{3.10}$$

In principle, the positive capacitance C and the inductance L can be adjusted to whatever values one wishes.

Such a measuring system can make arbitrarily quick and accurate measurements of $\hat{p}$, and of the classical force F which drives $\hat{p}$. One way to see this is by a semiclassical voltage-amplifier analysis of the type sketched in Eqs. (3.6)–(3.8). Another way is by a fully quantum-mechanical analysis corresponding to the case of a charge or current amplifier in series with the circuit (which now has $C=\infty$), rather than a voltage amplifier in parallel. In this analysis we leave the amplifier out of the circuit initially; we let the circuit evolve freely until a reasonably strong current is flowing; and we then insert our amplifier and quickly measure that current, or measure the charge on the infinite capacitor C. The free evolution of the system is governed by the Heisenberg equations for Hamiltonian (3.10):

$$\frac{d\hat{x}}{dt}=\frac{\hat{p}}{m}+K\hat{Q},\quad \frac{d\hat{p}}{dt}=F\,,$$
$$\frac{d\hat{Q}}{dt}=\frac{\hat{\Pi}}{L},\quad \frac{d\hat{\Pi}}{dt}=-K\hat{p}\,. \tag{3.11}$$

These Heisenberg equations are easily integrated to give

$$\begin{aligned}\hat{p}(t)&=\hat{p}(0)+Ft\,,\\ \hat{\Pi}(t)&=\hat{\Pi}(0)-K[\hat{p}(0)t+\tfrac{1}{2}Ft^2]\,,\\ \hat{Q}(t)&=\hat{Q}(0)+(1/L)[\hat{\Pi}(0)t-\tfrac{1}{2}K\hat{p}(0)t^2-\tfrac{1}{6}KFt^3]\,,\\ \hat{x}(t)&=\hat{x}(0)+\hat{p}(0)t+\tfrac{1}{2}Ft^2+K\hat{Q}(0)t\\ &\quad+(K/L)[\tfrac{1}{2}\hat{\Pi}(0)t^2-\tfrac{1}{6}K\hat{p}(0)t^3-\tfrac{1}{24}KFt^4]\,.\end{aligned} \tag{3.12}$$

From these integrals we can infer the following. If

the circuit is initially (at $t=0$) prepared in a Gaussian wave-packet state with

$$\langle\hat{\Pi}(0)\rangle=\langle\hat{Q}(0)\rangle=0\,,\qquad \Delta\Pi(0)=(L\hbar/2\tau)^{1/2},\quad \Delta Q(0)=(\hbar\tau/2L)^{1/2}\,, \tag{3.13}$$

and if the "free mass" is initially in a near-eigenstate of $\hat{p}$ with eigenvalue p_0, then after time τ has elapsed the expectation values and variances of the circuit variables are

$$\langle\hat{\Pi}(\tau)\rangle=-K(p_0\tau+\tfrac{1}{2}F\tau^2),\quad \Delta\Pi(\tau)=(L\hbar/2\tau)^{1/2}\,,$$
$$\langle\hat{Q}(\tau)\rangle=-(K/L)(\tfrac{1}{2}p_0\tau^2+\tfrac{1}{6}F\tau^3),\quad \Delta Q(\tau)=(\hbar\tau/L)^{1/2}\,. \tag{3.14}$$

At time τ we go into the circuit, disconnect it from the transducer if we wish, and measure either $\hat{\Pi}$ (the flux in the inductor), or $\hat{\Pi}/L$ (the current in the circuit), or $\hat{Q}$ (the charge on the infinite capacitor). With appropriately designed amplifiers, in principle we can make one or another of these measurements to within the variances (3.14), in a time $\lesssim\tau$. [This can be verified using the standard quantum limit on the noise performances of amplifiers. Note, moreover, that the precisions desired, $\delta\Pi\sim(L\hbar/2\tau)^{1/2}$ and $\delta Q\sim(\hbar\tau/L)^{1/2}$, are sufficiently modest that the uncertainty principle $\delta\Pi\delta Q\geq\frac{1}{2}\hbar$ even permits us to make the $\hat{\Pi}$ and $\hat{Q}$ measurements simultaneously!] From the measured value of $\hat{\Pi}$ or $\hat{\Pi}/L$ or $\hat{Q}$ we can infer p_0, in the absence of an external force F, to within probable error

$$\delta p_0\sim\frac{\Delta Q(\tau)}{\partial\langle\hat{Q}(\tau)\rangle/\partial p_0}\sim\frac{\Delta\Pi(\tau)}{\partial\langle\hat{\Pi}(\tau)\rangle/\partial p_0}$$
$$\sim\left(\frac{L\hbar}{K^2\tau^3}\right)^{1/2}\sim\left(\frac{m\hbar}{\tau}\right)^{1/2}\beta^{-1/2} \tag{3.15a}$$

or, if p_0 is known to this precision from previous measurements, we can infer the external force F to within probable error

$$\delta F\sim(L\hbar/K^2\tau^5)^{1/2}\sim(m\hbar/\tau^3)^{1/2}\beta^{-1/2}\,. \tag{3.15b}$$

Here

$$\beta=\frac{\text{energy in circuit}}{\text{energy of free mass}}$$
$$=\frac{\langle\hat{\Pi}(\tau)\rangle^2/2L}{p_0^2/2m}=\frac{K^2m\tau^2}{L} \tag{3.15c}$$

is a Gibbons-Hawking- (1971) type dimensionless coupling constant that will be discussed in Paper II. Equations (3.15) reveal that, no matter how quick (τ) the entire experiment must be, we can make the coupling constant K (or β) large enough in principle to produce any desired accuracy for our inferred values of the "free-mass" momentum p_0 and force F.

The above argument is similar to the one by which Ahanonov and Bohm (1961, 1964) refute a common misinterpretation of the energy-time uncertainty relation; cf. footnote 6. The Aharonov-Bohm argument has been criticized by Fock (1962) because it involves turning the coupling constant K on and off at $t=0$ and $t=\tau$, so that the mass m will be truly free of all coupling before and after the experiment. [It should be emphasized that, for the design of Fig. 6(a), one cannot turn K on and off merely by turning the magnetic field on and off, because doing so induces a position coupling which perturbs $\hat{p}$. In addition, one must modify Fig. 6(a) by adding a position coupling which is designed to cancel precisely the induced position coupling.] Fock suspects that the turn-on and turn-off cannot be done with the required precision. We, like Aharonov and Bohm (1964), disagree with Fock—but Fock's worries and our disagreement are irrelevant to the present analysis, because our objective here is merely to measure the momentum p_0 and force F with arbitrary accuracy, and that can be done without any time changes in the coupling constant K.

B. Measurements of a harmonic oscillator

We now return to our discussion of harmonic oscillators, and present a *Gedanken* experiment which shows that the $\hat{X}_1$ of an oscillator can be measured arbitrarily quickly and accurately, in principle. Here we shall describe our *Gedanken* experiment in somewhat abstract terms—focusing attention on the dynamical variables of the system and measuring apparatus, on the Hamiltonian which governs their evolution, and on a mathematical sketch of the measurement process and its potential accuracy. In Appendix B we describe apparatus which, in principle, could give a physical realization of the experiment; and in Appendix C we present a full mathematical analysis of the measurement process, complete with "reduction of the wave function" and repetitive measurements.

The oscillator to be measured is described by the variables of Eqs. (2.1)–(2.5), including coordinate $\hat{x}$, momentum $\hat{p}$, complex amplitude $\hat{X}_1+i\hat{X}_2$, frequency ω, and mass m.

The measuring apparatus consists of three parts: a "generator," which provides energy for and regulates the sinusoidal coupling of the interaction Hamiltonian; a "meter," which is coupled to $\hat{X}_1$ by the generator; and a "readout system" for studying the $\hat{X}_1$-induced motion of the meter.

The generator is a quantum-mechanical oscillator, which has precisely the same frequency ω as the oscillator being measured. Before the measurement, the generator is prepared in a coherent state of arbitrarily large amplitude. As is discussed in Appendix B.1.c, this means that the generator can be treated classically, and that it is not loaded by the experimental apparatus—and, consequently, that it produces perfect "$\cos\omega t$" and "$\sin\omega t$" terms in the Hamiltonian.

The meter is a one-dimensional quantum-mechanical "free mass," with generalized coordinate $\hat{Q}$, generalized momentum $\hat{\Pi}$, and generalized mass L. The coupling of the meter to the oscillator's $\hat{X}_1$ will be so strong that a tiny change δX_1 will make Q "swing" by an amount large compared to the width of its wave packet [cf. Eqs. (3.19)]. This swinging can then be observed with a classical readout system—i.e., we can put the "quantum-classical cut" of our analysis between the meter and the readout system, thereby avoiding the necessity to describe the readout system quantum mechanically; see discussion in Appendix C.

The total Hamiltonian for the coupled system, excluding the readout, is

$$\hat{H}=\hat{H}_0+\hat{H}_M+\hat{H}_I\,, \tag{3.16a}$$
$$\hat{H}_0=\hat{p}^2/2m+\tfrac{1}{2}m\omega^2\hat{x}^2\,, \tag{3.16b}$$
$$\hat{H}_M=\hat{\Pi}^2/2L\,, \tag{3.16c}$$
$$\hat{H}_I=K[\hat{x}\cos\omega t-(\hat{p}/m\omega)\sin\omega t]\hat{Q}=K\hat{X}_1\hat{Q}\,. \tag{3.16d}$$

Here $\hat{H}_0$ is the Hamiltonian of the free oscillator, $\hat{H}_M$ is the Hamiltonian of the meter,[7] and $\hat{H}_I$ is the interaction Hamiltonian for the oscillator coupled, via the classical generator ($\cos\omega t$ and $\sin\omega t$ terms), to the meter. In the Heisenberg picture of quantum mechanics the state of the system is constant, and the observables $\hat{X}_1 = \hat{x}\cos\omega t - (\hat{p}/m\omega)\sin\omega t$, $\hat{X}_2 = \hat{x}\sin\omega t + (\hat{p}/m\omega)\cos\omega t$, $\hat{Q}$, and $\hat{\Pi}$ evolve in accordance with the Heisenberg equations

$$d\hat{X}_1/dt = 0\,, \tag{3.17a}$$

$$d\hat{X}_2/dt = -(K/m\omega)\hat{Q}\,, \tag{3.17b}$$

$$d\hat{Q}/dt = \hat{\Pi}/L\,, \tag{3.17c}$$

$$d\hat{\Pi}/dt = -K\hat{X}_1\,. \tag{3.17d}$$

These are algebraically identical to the classical Hamiltonian equations of the system. Note that $\hat{X}_1$ is completely unaffected by the coupling to the measuring apparatus.

We presume that at time $t = t_0$, before the measurement begins, the oscillator is in a state (perhaps pure; perhaps mixed) with probability distribution $\mathcal{P}(X_1)$ whose expectation value is $\langle\hat{X}_1(t_0)\rangle = \xi_0$ and whose variance is $\Delta X_1(t_0) = \Sigma$. At $t = t_0$ the meter is prepared in a pure Gaussian wave-packet state with $\langle\hat{Q}(t_0)\rangle = \langle\hat{\Pi}(t_0)\rangle = 0$, $\Delta Q(t_0) = (\hbar\tau/2L)^{1/2}$, $\Delta\Pi(t_0) = (\hbar L/2\tau)^{1/2}$, where τ is the duration of the planned measurement. These variances are chosen to minimize the variance of $\hat{Q}$ after a time τ. The preparation of the meter can be done either with the oscillator-meter coupling turned on (in which case K is a constant before, during, and after the entire experiment), or with the coupling turned off ($K = 0$ for $t < t_0$, $K = \text{const}$ for $t_0 < t < t_0 + \tau$). The probability distribution $\mathcal{P}(X_1)$ is left unaffected by the physical manipulations of Q and Π involved in the preparation; cf. Eq. (3.17a).

After preparation of the meter, the system is allowed to evolve freely [Eqs. (3.17)] for a time τ. During this interval $\hat{X}_1$ is conserved, and the evolution of $\hat{Q}$ is given by

$$\hat{Q}(t) = \hat{Q}(t_0) + \frac{\hat{\Pi}(t_0)}{L}(t - t_0) - \frac{K\hat{X}_1}{2L}(t - t_0)^2\,. \tag{3.18}$$

The interaction produces a strong correlation between the states of the oscillator and meter: At time $t_1 = t_0 + \tau$ the meter's mean coordinate gets displaced to

$$\langle\hat{Q}(t_1)\rangle = -(K\tau^2/2L)\xi_0 \tag{3.19a}$$

[cf. Eq. (3.18)], while its variance grows to

$$\Delta Q(t_1) = [(\hbar\tau/L) + (K\tau^2/2L)^2\Sigma^2]^{1/2}\,. \tag{3.19b}$$

At time t_1 the readout system "reads out" a value Q_m for the meter coordinate, where Q_m is likely to lie somewhere in the range $\langle\hat{Q}(t_1)\rangle \pm$ (a few) $\times\,\Delta Q(t_1)$. (This readout can be done leaving the coupling K on, or turning it off, as one wishes; it makes no difference.) Using formula (3.19a) the experimenter infers from Q_m a value

$$\xi_m = -(2L/K\tau^2)Q_m \tag{3.20a}$$

for X_1. In a set of measurements on an ensemble of identical systems, the mean of this inferred value is ξ_0, and its variance is

$$\Delta\xi_m = (2L/K\tau^2)\Delta Q(t_1) = (\Sigma^2 + 4\hbar L/K^2\tau^3)^{1/2}\,. \tag{3.20b}$$

Of course, as Eq. (3.20b) shows, one cannot determine ξ_0 accurately if the probability distribution $\mathcal{P}(X_1)$ has a large spread; however, if $\mathcal{P}(X_1)$ is highly peaked about ξ_0 [$\Sigma \ll (4\hbar L/K^2\tau^3)^{1/2}$], the measurement can determine ξ_0 with a probable error

$$\Delta\xi_m \simeq (4\hbar L/K^2\tau^3)^{1/2} \sim (\hbar/2m\omega)^{1/2}(\beta\omega\tau)^{-1/2}\,. \tag{3.21a}$$

Here

$$\beta = \frac{\text{energy in circuit}}{\text{energy of } X_1 \text{ motion of oscillator}} = \frac{\frac{1}{2}L\langle\dot{\hat{Q}}(t_1)\rangle^2}{\frac{1}{2}m\omega^2\xi_0^2} = \frac{K^2\tau^2}{Lm\omega^2} \tag{3.21b}$$

is a dimensionless coupling constant. No matter how small the measurement time τ may be, by choosing K^2/L large enough (i.e., β large enough) one can make the measurement error $\Delta\xi_m$ as small as one wishes. [Note that this remains true even if the readout of Q is much less accurate than $(\hbar\tau/L)^{1/2}$; see analysis in Appendix C.7.] The measurements can be "arbitrarily quick and arbitrarily accurate."

If a weak, classical force is driving the oscillator [term $-\hat{x}F(t)$ added to the Hamiltonian; cf. Eq. (2.18)], then during the time τ the expectation value of $\hat{X}_1$ changes by an amount

$$\delta\xi = -\int_{t_0}^{t_1}\frac{F(t')}{m\omega}\sin\omega t'\,dt' \tag{3.22}$$

[cf. Eq. (2.31)]; and the meter's mean coordinate gets displaced to

$$\langle\hat{Q}(t_1)\rangle = -(K\tau^2/2L)(\xi_0 + \mathfrak{F})\,, \tag{3.23a}$$

$$\mathfrak{F} \equiv -\frac{2}{\tau^2}\int_{t_0}^{t_1}dt'\int_{t_0}^{t'}dt''\int_{t_0}^{t''}dt'''\,\frac{F(t''')}{m\omega}\sin\omega t'''\,, \tag{3.23b}$$

while the variance of $\hat{Q}(t_1)$ is still given by Eq. (3.19b). For measurement times short enough that $F(t)\sin\omega t$ is nearly constant during the measurement

$$\mathfrak{F} \simeq \tfrac{1}{3}\,\delta\xi \simeq -\tfrac{1}{3}\tau[F(t_0)/m\omega]\sin\omega t_0\,.$$

If ξ_0 is known from previous measurements to within the error (3.21), a measurement of $\hat{Q}$ at time t_1 allows one to determine $\mathfrak{F}$ (or $\delta\xi$) with accuracy

$$\Delta\mathfrak{F} \sim (\hbar L/K^2\tau^3)^{1/2}\,. \tag{3.24}$$

Such a measurement permits one (in principle) to monitor the force F arbitrarily quickly and accurately, in the limit as τ and $(\hbar L/K^2\tau^5)^{1/2}$ are made arbitrarily small.

The preceding analysis is rigorous, but it is far from complete. In Appendix C we present a more detailed analysis; in particular, we analyze a *sequence* of measurements of X_1, including the "reduction of the wave function" at the end of each measurement. This more detailed analysis allows us to investigate the behavior of X_1 and X_2 during a sequence of measure-

[7] To achieve a Hamiltonian of the form of Eq. (3.16), the measuring systems described in Appendix B must incorporate a negative capacitor or a negative spring, which converts a velocity sensor into a momentum sensor (cf. Sec. III.A.2). For these systems the *free* meter is a "mass on a negative spring"; the coupling to the oscillator converts the meter into a "free mass."

ments. The most important results concern X_2. We show that a "feedback force" on the meter can keep the expectation value of $\hat{X}_2$ close to zero. However, the variance of $\hat{X}_2$ inevitably increases during a sequence, and the increase is proportional to the square root of the number of measurements. Practical implications of this "random walk of X_2" are discussed in Paper II.

IV. FORMAL DISCUSSION OF QUANTUM NONDEMOLITION MEASUREMENT[8]

Here we shall distill from our analysis of oscillators and free masses the essence of "quantum nondemolition measurement" and label that essence using the formal and precise language of nonrelativistic quantum mechanics. The final product may be unpalatable to the practical-minded reader, but we hope it will clarify the fundamental principles underlying "nondemolition measurement."

A. Definition of quantum nondemolition measurement and its implications

Our investigation of quantum nondemolition measurement was stimulated by the desire to monitor a classical force acting on a harmonic oscillator with better accuracy than can be obtained using standard "amplitude-and-phase" techniques. Braginsky (1970), and later Giffard (1976), had pointed out the limitations of the standard techniques (see Sec. II.C), and Braginsky and Vorontsov (1974) had proposed overcoming these limitations by making what they called "quantum nondemolition measurements." In such a measurement one monitors a single observable of the oscillator, and it must be an observable that can be measured over and over again with the result of each measurement being completely determined (in the absence of a classical force) by the result of an initial, precise measurement. The force is detected by changes it produces in this sequence of precisely predictable values.

The key feature of such a nondemolition measurement is *repeatability*—once is not enough! If one can couple strongly enough to a physical system, then any of its observables can be measured (in principle) with arbitrary precision at a particular instant. (This is the content of a controversial general "theorem" in nonrelativistic quantum theory; see discussion in footnote 6). Such a precise measurement "localizes" the system at the measured value of the observable. An initial, precise measurement can be regarded as preparing the system in a state with a nearly definite value of the measured observable. The goal of a subsequent measurement is to determine this value. However, the initial, precise measurement inevitably produces huge uncertainties in observables that do not commute with the measured observable, and in general, these uncertainties "feed back" into the measured observable as the system evolves. Consequently, the result of a subsequent measurement is uncertain. If one wishes to make *repeated* precise measurements whose results are completely predictable (no uncertainty!), one must measure an observable that does not become contaminated by uncertainties in other, noncommuting observables.

It is easy to formulate a general condition for making such a sequence of completely predictable measurements: The system being measured must be in an eigenstate of the measured observable at the time of each measurement. Then the result of each measurement is exactly equal to the eigenvalue at the time of the measurement, and immediately after the measurement the system is left in the same eigenstate ("measurement of the first kind"; cf. footnote 6). This condition clarifies the nature of nondemolition measurement and, at the same time, makes it clear that what is not being demolished is the state of the system; it is left unchanged by each measurement except for an unknown (and irrelevant) phase factor.

To formalize these ideas, consider an arbitrary quantum-mechanical system with free Hamiltonian $\hat{H}_0$. The objective is to measure an observable $\hat{A}$ of this system. ($\hat{A}$ is a Hermitian operator; it may have *explicit* time dependence.) For a resonant-bar gravitational-wave detector, the system would be the fundamental mode of the bar, which can be idealized as a simple harmonic oscillator; and $\hat{A}$ might be the number of quanta in the oscillator or the real part of the oscillator's complex amplitude. For such a detector, one measures $\hat{A}$ in order to monitor the classical force on the oscillator produced by a gravitational wave; to allow for that possibility here, we include in the Hamiltonian a term $\hat{D}F(t)$, where $\hat{D}$ is some observable of the system and $F(t)$ is the "classical force." To ensure that $\hat{A}$ responds to $F(t)$, we require $[\hat{A},\hat{D}]\neq 0$.

In order to measure $\hat{A}$, one must couple the system to a measuring apparatus. The details of the system's interaction with the measuring apparatus are described by the interaction Hamiltonian $\hat{H}_I$, which contains all terms in the Hamiltonian that depend on variables of both the system and the measuring apparatus. The total Hamiltonian—including the system, its coupling to the "classical force," and the measuring apparatus—has the form

$$\hat{H}=\hat{H}_0+\hat{D}F(t)+\hat{H}_I+\hat{H}_M, \tag{4.1}$$

where $\hat{H}_M$ is the Hamiltonian of the measuring apparatus—i.e., that part of the Hamiltonian which depends only on measuring apparatus variables [cf. Eq. (2.36)].

We now define a *quantum nondemolition* (QND) *measurement* of $\hat{A}$ as a sequence of precise measurements of $\hat{A}$ such that the result of each measurement (after the first) is completely predictable (in the absence of a classical force) from the result of the preceding measurement. If an observable can be measured this way (in principle), we call it a *quantum nondemolition observable.*

This definition can be used to derive a condition for a QND observable—a condition most easily formulated by ignoring the details of the interaction with the measuring apparatus. This is not to say that these details are unimportant: For example, the strength of the coupling between the system and measuring apparatus determines how quickly a given measurement precision

[8]The ideas and prose of this section are due entirely to Carlton M. Caves, and constitute a portion of the material submitted by him to the California Institute of Technology in partial fulfillment of the requirements for the Ph.D. degree.

can be achieved (see Sec. III). However, the *fundamental* limits on the predictability of a sequence of measurements of $\hat{A}$ are determined not by the interaction with the measuring apparatus, but by uncertainties (variances of observables) which are built into a quantum-mechanical description of the free evolution of the system. Of course, the interaction with the measuring apparatus, if chosen poorly, can ruin a QND measurement by increasing the variance of the measured observable; however, as we demonstrate in Sec. IV. B, the interaction need not degrade the measurement at all in principle. Therefore, for the remainder of this subsection, we ignore the interaction term in the Hamiltonian; we simply assume that there is a way to measure $\hat{A}$ with arbitrary precision at any instant (infinitely strong coupling!) and that such a measurement leaves the system in an eigenstate of $\hat{A}$ whose eigenvalue is the measured value ("measurement of the first kind"; cf. footnote 6). We also ignore, for the moment, the classical force.

We now consider a sequence of measurements of $\hat{A}$. The analysis proceeds most smoothly in the Heisenberg picture of quantum mechanics, which we use throughout the rest of this subsection. The initial measurement is made at time t_0, and we assume that the experimenter has no control over the state of the system before this initial measurement. (This may be a bad assumption; see discussion in Sec. IV. C.) The normalized eigenstates of $\hat{A}(t_0)$ are denoted by $|A,\alpha\rangle$, where $\hat{A}(t_0)|A,\alpha\rangle = A|A,\alpha\rangle$ and where α labels the states in any degenerate subspaces of $\hat{A}(t_0)$.

The result of the initial measurement is one of the eigenvalues A_0 of $\hat{A}(t_0)$, and the state of the system immediately after the measurement is an eigenstate of $\hat{A}(t_0)$ with this eigenvalue: $|\Psi(t_0)\rangle = \sum_\alpha c_\alpha |A_0,\alpha\rangle$, where the c_α's are arbitrary (subject to normalization) constants. In the interval before the next measurement the system evolves freely, and in the Heisenberg picture the state of the system does not change: $|\Psi(t)\rangle = |\Psi(t_0)\rangle$. If a second measurement at $t=t_1$ is to yield a completely predictable result, then all of the states $|A_0,\alpha\rangle$ must be eigenstates of $\hat{A}(t_1)$ with the same eigenvalue, although the new eigenvalue need not equal A_0. Hence, one obtains the requirement

$$\hat{A}(t_1)|A_0,\alpha\rangle = f_1(A_0)|A_0,\alpha\rangle \text{ for all } \alpha, \tag{4.2}$$

where f_1 is an arbitrary real-valued function. Equation (4.2) guarantees that the result of a measurement at $t=t_1$ will be $f_1(A_0)$, because $|\Psi(t_0)\rangle$ will be an eigenstate of $\hat{A}(t_1)$ with eigenvalue $f_1(A_0)$ for arbitrary c_α's.

By assumption, the result of the initial measurement can be any of the eigenvalues of $\hat{A}(t_0)$. Thus Eq. (4.2) must hold for all values of A_0, and $\hat{A}(t_1)$ must satisfy the operator equation $\hat{A}(t_1) = f_1[\hat{A}(t_0)]$.[9] In a sequence of measurements a similar operator equation must hold at each step in the sequence. Therefore one obtains the following set of requirements for a QND observable that is to be measured at times $t=t_0,\ldots,t_n$:

$$\hat{A}(t_k) = f_k[\hat{A}(t_0)] \text{ for } k=1,\ldots,n, \tag{4.3}$$

where each f_k is some real-valued function. These formal constraints on the free evolution of $\hat{A}$ in the Heisenberg picture embody the fundamental principle of QND measurement: If the system begins in an eigenstate of $\hat{A}$, its free evolution must leave it in an eigenstate of $\hat{A}$ at the time of each measurement. The conditions (4.3) for a QND observable were given previously by the authors (Thorne *et al.*, 1978).

One is usually interested in making QND measurements at arbitrary times or continuously. Then Eq. (4.3) must hold for all times:

$$\hat{A}(t) = f[\hat{A}(t_0);t,t_0]. \tag{4.4}$$

An observable that satisfies Eq. (4.4) we call a *continuous* QND observable. An observable that satisfies Eq. (4.3) only at carefully selected times we call a *stroboscopic* QND observable. Examples of stroboscopic QND observables are the position and momentum of a harmonic oscillator (stroboscopic measurement; see Sec. II. F. 2). Because of their importance, we restrict our attention to continuous QND observables for the rest of this section.

The simplest way to satisfy Eq. (4.4) is to choose an observable which is conserved in the absence of interactions with the external world:

$$0 = \frac{d\hat{A}}{dt} = -\frac{i}{\hbar}[\hat{A},\hat{H}_0] + \frac{\partial \hat{A}}{\partial t}. \tag{4.5}$$

For example, the continuous QND observables we have considered for a harmonic oscillator—$\hat{X}_1$, $\hat{X}_2$, and $\hat{N}$—are conserved. Note that the free Hamiltonian $\hat{H}_0$ is always a QND observable (provided $\partial \hat{H}_0/\partial t = 0$).

It is harder to find nontrivial examples of nonconserved continuous QND observables. One system which has such observables is a mass m on a "negative spring"—i. e., a mass with Hamiltonian $\hat{H}_0 = \hat{p}^2/2m - \frac{1}{2}m\omega^2\hat{x}^2$. For such a system the observables $\hat{x} \pm (\hat{p}/m\omega)$ are QND observables, but they are not conserved.

It is useful to note here an important commutation property satisfied by any continuous QND observable $\hat{A}$:

$$[\hat{A}(t),\hat{A}(t')] = 0 \text{ for all times } t \text{ and } t'. \tag{4.6}$$

This property follows immediately from the QND condition (4.4). Equivalent to Eq. (4.6) is the statement that $\hat{A}$ commutes with all its derivatives—i. e.,

$$\begin{aligned} 0 &= \left[\hat{A}, \frac{d^n\hat{A}}{dt^n}\right] \\ &= \left[\hat{A}, \sum_{l=0}^{n}\left(-\frac{i}{\hbar}\right)^{n-l}\binom{n}{l}\left[\frac{\partial^l\hat{A}}{\partial t^l}, \hat{H}_0\right]^{(n-l)}\right] \\ &\text{for } n=1,2,3,\ldots, \end{aligned} \tag{4.7}$$

where

$$[\hat{C},\hat{D}]^{(n)} \equiv \begin{cases} \hat{C}, & n=0, \\ \Big[\Big[\ldots[[\hat{C},\underbrace{\hat{D}],\hat{D}],\ldots],\hat{D}}_{n\ \hat{D}\text{'s}}\Big], & n=1,2,3,\ldots. \end{cases} \tag{4.8}$$

The latter equality in Eq. (4.7) can be obtained (provided $\partial \hat{H}_0/\partial t = 0$) by using the operator equations of

[9] One might wish to require that there exist a one-to-one correspondence between the possible measured values at $t=t_0$ and $t=t_1$, in which case f_1 must be invertible.

motion in the Heisenberg picture.

Unruh (1979) has recently considered the problem of nondemolition measurement. He discusses many of the issues considered in this section, but from a somewhat different point of view. He has proposed that Eq. (4.7) [or, equivalently, Eq. (4.6)] be used to characterize QND observables. {Actually, Unruh considers only observables with no explicit time dependence—a serious restriction which rules out such very important observables as the $\hat{X}_1$ of an oscillator. Because of this restriction, Unruh's quantum nondemolition condition is

$$[\hat{A},[\hat{A},\hat{H}_0]^{(n)}]=0 \quad \text{for } n=1,2,3,\ldots\,, \tag{4.7'}$$

which is the specialization of Eq. (4.7) to the case $\partial\hat{A}/\partial t=0$.}

The motivation for Unruh's definition is discussed in Sec. IV.B, but for now it is important to note that, although Eq. (4.7) is an immediate consequence of the QND condition (4.4), the implication cannot be reversed. The observables satisfying Eq. (4.7) constitute a more general class than the QND observables; we call such observables *generalized* (continuous) QND *observables*.

Examples of observables which satisfy Eq. (4.7), but not Eq. (4.4), can be obtained by considering a system suggested by Unruh (1979): a charged particle (charge e, mass m) moving in a uniform magnetic field $\mathbf{B}=B_0\mathbf{e}_z$ and an electric field $\mathbf{E}=(eB_0^2/8m)\nabla(x^2+y^2-2z^2)$. If the vector potential is chosen to be $\mathbf{A}=\frac{1}{2}B_0(-y\mathbf{e}_x+x\mathbf{e}_y)$, then $\hat{p}_x$ and $\hat{p}_y$ (x and y components of the particle's canonical momentum) are generalized QND observables, but they do not satisfy the QND criterion (4.4). [For this system the observables $\hat{x}-(\hat{p}_x/m)t$ and $\hat{y}-(\hat{p}_y/m)t$ are also generalized QND observables.]

Any generalized QND observable $\hat{A}$ does obey an evolution constraint similar to the QND constraint (4.4). Successive differentiation of Eq. (4.7) shows that all derivatives of $\hat{A}$ mutually commute, and a Taylor expansion of $\hat{A}$ about some initial time t_0 shows that the free evolution of $\hat{A}$ must have the form

$$\hat{A}(t)=f[\hat{A}(t_0);\hat{B}_1,\ldots,\hat{B}_n;t,t_0], \tag{4.9}$$

where the Hermitian operators $\hat{B}_i$ commute with one another and with $\hat{A}(t_0)$. In writing Eq. (4.9), it is assumed that none of the operators $\hat{B}_i$ can be written as a function of $\hat{A}(t_0)$ and the other $\hat{B}_i$'s; otherwise, the functional dependence of $\hat{A}(t)$ could be simplified. Note that if $\hat{A}(t_0)$ has no degeneracies, the only operators which commute with it are functions of itself; hence, a nondegenerate generalized QND observable is automatically a QND observable.

Generalized QND observables can be compared most tellingly with QND observables by using Eq. (4.9). The key difference is the following: A system which begins in any eigenstate of a QND observable remains in an eigenstate of that observable; this is true for a generalized QND observable only if the initial eigenstate is a simultaneous eigenstate of $\hat{A}(t_0)$ and the $\hat{B}_i$'s. An equivalent manifestation of this difference is that the result of a given measurement of a generalized QND observable cannot be predicted solely from the result of one preceding measurement of that observable. However, it can be predicted from the results of several preceding measurements of $\hat{A}$—enough to specify the values of $\hat{A}(t_0)$ and each of the $\hat{B}_i$'s; alternatively, it can be predicted from a simultaneous measurement of $\hat{A}(t_0)$ and all the $\hat{B}_i$'s.

It should be clear that generalized QND observables provide the key to extending the concept of quantum nondemolition to *sets* of observables. Such a set of QND observables would have the following properties: (i) the observables could be measured simultaneously with arbitrary precision; and (ii) in a sequence of precise, simultaneous measurements of all the observables, the results of each set of measurements could be predicted from the results of the preceding set. A set of QND observables $\hat{A}_i$ would obey the commutation constraints $[\hat{A}_i(t),\hat{A}_j(t')]=0$ for all observables in the set and for all times t and t'. In the above charged-particle example, $\hat{p}_x$ and $\hat{p}_y$ form a set of QND observables.

Having defined QND measurement, we now consider its application to the problem of monitoring a classical force $F(t)$. The procedure for monitoring $F(t)$ is to make a sequence of measurements of a QND observable and to detect the force by the changes it produces in the precisely predictable values which would be measured in the absence of the force.

One would like to do more than simply "detect" the force: Ideally, one would like to monitor its time dependence with arbitrary accuracy; and if the force is arbitrarily classical, there is no reason in principle why one cannot do so. In fact, a sequence of measurements of the observable $\hat{A}$ can reveal with arbitrary accuracy the time evolution of $F(t)$ if and only if the following conditions are satisfied: (i) The measuring apparatus and its coupling to the measured system [$\hat{H}_M$ and $\hat{H}_I$ of Eq. (4.1)] must be chosen so as to produce instantaneous and arbitrarily precise measurements of $\hat{A}$ (see Sec. IV.B). (ii) The measurements must be made at arbitrarily closely spaced times. (iii) The result of the $(k+1)$th measurement at time t_k must be uniquely determined by the result of the initial measurement at time t_0 plus the time history $F(t')$ of the force between t_0 and t_k. This is possible if and only if $\hat{A}$ is a continuous QND observable in the presence of the driving force F [Eq. (4.4)]:

$$\hat{A}(t)=f[\hat{A}(t_0);F(t');t,t_0] \quad \text{for } t_0<t'<t\,, \tag{4.10a}$$

where $\hat{A}(t)$ is the Heisenberg-picture evolution of $\hat{A}$ under the action of $\hat{H}=\hat{H}_0+\hat{D}F$. Here f is a function of $\hat{A}(t_0)$, t, and t_0, and is a functional of $F(t')$. (iv) From the time history of the measured values of $\hat{A}(t)$ one must be able to compute uniquely the time history of $F(t)$. The measured values will be

$$A(t)=f[A_0;F(t');t,t_0],$$

where A_0 is the (arbitrary) eigenvalue of $\hat{A}(t_0)$ obtained in the first measurement. Thus condition (iv) is equivalent to the demand that

$$A(t)\equiv f[A_0;F(t');t,t_0] \text{ must be a uniquely invertible functional of } F(t'), \text{ for every eigenvalue } A_0 \text{ that is a possible result of the first measurement of } \hat{A}(t_0). \tag{4.10b}$$

Of these conditions only Eqs. (4.10a) and (4.10b) are constraints on the choice $\hat{A}$ of the observable to be measured. Thus, for a given system and a given coupling to the classical force F (i.e., for given $\hat{H}=\hat{H}_0+\hat{D}F$), conditions (4.10) are necessary and sufficient to permit in principle measurements of $\hat{A}$ that reveal with arbitrary accuracy the time evolution of $F(t)$. To such an observable $\hat{A}$ we shall give the name "QNDF." Because a QNDF observable is QND in the presence of the force F, it will necessarily satisfy Eq. (4.6)—$[\hat{A}(t),\hat{A}(t')]=0$—and also the first equality of Eq. (4.7). These same two equations are also satisfied by "generalized QNDF observables"—i.e., observables for which the functional f of Eqs. (4.10) depends also on a set of mutually commuting Hermitian operators $\hat{B}_i$ which all commute with $\hat{A}(t_0)$.

The distinction between QND and QNDF observables arose earlier in comparing quantum-counting measurements and measurements of $\hat{X}_1$ as ways of monitoring a force acting on a harmonic oscillator (see Secs. II.D and II.E). Measurements of $\hat{X}_1$ can be used to monitor an arbitrarily weak force $F(t)$ with arbitrary accuracy, in principle; quantum counting can "detect" an arbitrarily weak force, but it cannot provide good accuracy in monitoring the force's precise time dependence. The fundamental reason for this difference is that $\hat{X}_1$ is a QNDF observable, while $\hat{N}$ is not.

In his recent treatment of nondemolition measurement Unruh (1979) has also drawn attention to the important distinction between QND and QNDF observables (QNDR and QNDD, respectively, in his notation).

B. Interaction with the measuring apparatus

Up until now we have neglected the details of the measuring apparatus which is actually used to measure a QND observable $\hat{A}$. We now rectify this omission. Our main concern is to demonstrate our earlier assertion that the interaction between the system and the measuring apparatus need not degrade the quality of a QND measurement at all, in principle. The analysis in this subsection is restricted to continuous observables, but it can easily be modified to handle stroboscopic observables.

In a real experiment the measuring apparatus consists of a series of components. Each component is coupled to the preceding component, and only the first stage in the series directly "contacts" the system. Fortunately, we need not concern ourselves with this entire complicated structure; its complexities can remain buried in the measuring apparatus Hamiltonian $\hat{H}_M$. We need only consider the first stage of the measuring apparatus and its interaction with the system—an interaction whose mathematical manifestation is the interaction Hamiltonian $\hat{H}_I$.

The measuring apparatus must actually respond to $\hat{A}$, and this demand means that $\hat{H}_I$ must depend on $\hat{A}$ and on one or more variables of the first stage of the measuring apparatus. In addition, the measuring apparatus ought not to respond to observables of the system other than $\hat{A}$, and this desire means that $\hat{A}$ ought to be the only observable of the system appearing in $\hat{H}_I$. The simplest interaction Hamiltonian of this form is

$$\hat{H}_I = K\hat{A}\hat{Q}\,, \tag{4.11}$$

where $\hat{Q}$ is some observable of the measuring apparatus and K is a coupling constant. This is the type of interaction Hamiltonian which was used in Secs. II.F.1 and III.B to analyze continuous two-transducer measurements of $\hat{X}_1$.

If $\hat{A}$ contains *explicit* time dependence, the coupling between the system and the measuring apparatus must be modulated so as to supply the proper time dependence in $\hat{H}_I$. The modulation must be provided by an external, *classical* "clock." Unruh (1979) has pointed out that any "clock" is an inherently quantum-mechanical device whose quantum properties cannot be ignored *a priori*; however, the "clock" can always be excited into a highly energetic, essentially classical state, where uncertainties due to its quantum-mechanical nature are unimportant. This issue is discussed in the context of measurements of $\hat{X}_1$ in Appendix B.1.c.

We now turn to the main concern of this subsection—to demonstrate the following fundamental property of QND observables. *The evolution of a continous* QND *observable* $\hat{A}$ *(in the Heisenberg picture) is completely unaffected by the interaction with the measuring apparatus (in the absence of a classical force), provided that* $\hat{A}$ *is the only observable of the system which appears in the interaction Hamiltonian.*[10] The proof of this property relies on only one feature of $\hat{A}$—that it satisfies Eq. (4.6) in the absence of the interaction with the measuring apparatus. Thus the property holds for generalized QND observables, and for QNDF observables even in the presence of the classical force.

Proving the property is not difficult, but it is sufficiently important that it is worthwhile to sketch the proof in some detail. We consider the case of a QND observable in the absence of a classical force, and we now let $\hat{A}$ denote the QND observable in the Schrödinger picture. The total Hamiltonian, now considered to be written in the Schrödinger picture, is given by

[10] The assumption in the text—that $\hat{A}$ is the only observable of the system which appears in the interaction Hamiltonian—is more restrictive than necessary. The argument given in the text actually proves the following more general theorem: The evolution of an observable $\hat{A}$ is unaffected by the interaction with the measuring apparatus if $[\hat{A}_I(t),\hat{\mathcal{H}}_I(t')]=0$ for all times t and t', where

$$\hat{A}_I(t)\equiv\hat{U}_0^{\dagger}(t,t_0)\hat{A}(t)\hat{U}_0(t,t_0)$$

and

$$\hat{\mathcal{H}}_I(t)\equiv\hat{U}_0^{\dagger}(t,t_0)\hat{U}_M^{\dagger}(t,t_0)\hat{H}_I(t)\hat{U}_M(t,t_0)\hat{U}_0(t,t_0)$$

are the interaction-picture forms of $\hat{A}$ and $\hat{H}_I$. (In the definitions of $\hat{A}_I$ and $\hat{\mathcal{H}}_I$, $\hat{A}$ and $\hat{H}_I$ are Schrödinger-picture observables.) This theorem allows one to loosen the text's assumption about the nature of $\hat{H}_I$. For example, if $\hat{A}$ is conserved in the absence of the interaction with the measuring apparatus, then it remains conserved in the presence of the interaction if $[\hat{A},\hat{H}_I]=0$. As a more general example, one can make the following statement about a "set of QND observables" (see Sec. IV.A): The evolution of each observable in the set is unaffected by the coupling to the measuring apparatus, provided that the only observables of the system which appear in $\hat{H}_I$ are members of the set.

Eq. (4.1) with the classical-force term deleted. We let $\hat{U}_0(t,t_0)$, $\hat{U}_M(t,t_0)$, and $\hat{U}(t,t_0)$ be the unitary time-development operators for $\hat{H}_0$, $\hat{H}_M$, and $\hat{H}$, respectively [cf. Eq. (2.19)]. The assumption about the nature of the interaction means that $\hat{H}_I$ has the form

$$\hat{H}_I = \hat{H}_I[\hat{A}(t); \hat{Q}_1, \ldots, \hat{Q}_n; t], \tag{4.12}$$

where the operators $\hat{Q}_i$ are observables of the measuring apparatus.

The two operators of interest are the interaction-picture and Heisenberg-picture forms of the QND observable:

$$\hat{A}_I(t) \equiv \hat{U}_0^\dagger(t,t_0)\hat{A}(t)\hat{U}_0(t,t_0), \tag{4.13a}$$

$$\hat{A}_H(t) \equiv \hat{U}^\dagger(t,t_0)\hat{A}(t)\hat{U}(t,t_0). \tag{4.13b}$$

The interaction-picture operator $\hat{A}_I(t)$ gives the evolution of the QND observable in the absence of the interaction with the measuring apparatus; thus it is the operator which satisfies the QND condition (4.4) and which, in particular, also satisfies Eq. (4.6). The Heisenberg-picture operator $\hat{A}_H(t)$ gives the evolution of the QND observable in the presence of the measuring apparatus. The object of the proof is to show that $\hat{A}_H(t) = \hat{A}_I(t)$.

The operators $\hat{A}_I(t)$ and $\hat{A}_H(t)$ are related by

$$\hat{A}_H(t) = \hat{\mathfrak{u}}^\dagger(t,t_0)\hat{A}_I(t)\hat{\mathfrak{u}}(t,t_0). \tag{4.14}$$

Here $\hat{\mathfrak{u}}(t,t_0) = \hat{U}_0^\dagger(t,t_0)\hat{U}_M^\dagger(t,t_0)\hat{U}(t,t_0)$ satisfies

$$i\hbar\frac{d\hat{\mathfrak{u}}(t,t_0)}{dt} = \hat{\mathcal{K}}_I(t)\hat{\mathfrak{u}}(t,t_0), \quad \hat{\mathfrak{u}}(t_0,t_0) = 1, \tag{4.15}$$

where

$$\hat{\mathcal{K}}_I(t) \equiv \hat{H}_I[\hat{A}_I(t); (\hat{Q}_1)_I(t), \ldots, (\hat{Q}_n)_I(t); t], \tag{4.16a}$$

$$(\hat{Q}_i)_I(t) \equiv \hat{U}_M^\dagger(t,t_0)\hat{Q}_i\hat{U}_M(t,t_0). \tag{4.16b}$$

The solution for $\hat{\mathfrak{u}}(t,t_0)$ can be written in the form

$$\hat{\mathfrak{u}}(t,t_0) = \mathbf{T}\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t}\hat{\mathcal{K}}_I(t')\,dt'\right), \tag{4.17}$$

where **T** means that all products are time ordered (see, e.g., Sec. 18.7 of Merzbacher, 1970). The fact that $\hat{A}_I(t)$ satisfies Eq. (4.6) guarantees $[\hat{A}_I(t), \hat{\mathfrak{u}}(t,t_0)] = 0$, which with Eq. (4.14) implies that $\hat{A}_H(t) = \hat{A}_I(t)$. As claimed, the QND observable is completely isolated from the measuring apparatus. A trivial extension of this argument proves the result for QNDF observables in the presence of the classical force.

The meaning of this fundamental property should be emphasized. The property says that the evolution of a QND observable, calculated using the equations of motion in the Heisenberg picture, is unaffected by interaction with the measuring apparatus. This means that the expectation value and variance of $\hat{A}$ evolve during a measurement exactly as they would have had the measuring apparatus been disconnected. "Noise" in the measuring apparatus does not feed back onto $\hat{A}$ and increase its variance. However, a complete description of a measurement requires more than just a calculation of the quantum-mechanical evolution: At some time the measurement must end, the quantum-mechanical evolution equations must be suspended, and the quantum state of the coupled system and measuring apparatus must be "reduced" to be consistent with the results of the measurement.

If the system begins the measurement in an eigenstate of $\hat{A}$, it remains in an eigenstate of $\hat{A}$ throughout the measurement, and the "reduction of the wave function" leaves it in the same eigenstate. This is an immediate consequence of the above fundamental property. However, in any real measurement the probability distribution of $\hat{A}$ has some variance, and at the time of "reduction" the expectation value of $\hat{A}$ "jumps" a distance which can be as large as the variance. In this sense the measuring apparatus does affect the QND observable. However, these "jumps" are a consequence of the fact that the measuring apparatus is not making absolutely precise measurements; they do not affect our conclusion that *in principle* the measuring apparatus need not degrade the predictability of a sequence of measurements of $\hat{A}$. For a detailed analysis of this issue in the context of measurements of $\hat{X}_1$, see Appendix C.

It is now clear why the details of the interaction with the measuring apparatus could be ignored in Sec. IV.A. There we assumed infinitely strong coupling so that precise measurements could be made instantaneously. For a realistic interaction, the coupling strength is finite, and a certain amount of time is required to achieve a desired measurement precision. However, no matter what the coupling strength may be and how long the measurement may last, a QND observable is completely unaffected by the coupling to the measuring apparatus if $\hat{H}_I$ has the required form. Indeed, for any measurement time one can achieve any desired accuracy by making the coupling strength large enough—i.e., the measurements can be arbitrarily quick and arbitrarily accurate. Of course, it may be difficult in practice to design an interaction which is sensitive only to $\hat{A}$; and if other observables of the system appear in $\hat{H}_I$, the time a measurement can take before it disturbs $\hat{A}$ significantly may be limited.

It is interesting to note here that if the right kind of interaction can be designed, a QND observable is isolated not only from "quantum noise" but also from "classical noise" in the measuring apparatus (thermal noise in resistors, shot noise in amplifiers, etc.). In this sense any QND measurement is a "back-action-evading" measurement, because the measured observable evades the back-action noise from the measuring apparatus.

As mentioned earlier, Unruh (1979) has proposed that Eq. (4.7′) be used to characterize nondemolition measurement. He considers only observables with no explicit time dependence, he assumes an interaction Hamiltonian of the form of Eq. (4.11), and he characterizes nondemolition measurement by the demand that the measured observable be completely isolated from the measuring apparatus. As we have shown, any generalized QND observable meets this demand. Thus it is not surprising that Unruh's QND condition is Eq. (4.7′)—the condition for a generalized QND observable with no explicit time dependence.

C. Comments and caveats

The discussion of nondemolition measurement in this section has been presented in the formal language of nonrelativistic quantum mechanics, and the description of the measurement process has been highly idealized. The reader can be forgiven for asking whether these idealized descriptions have anything to do with *real* experiments. We think so, and the best evidence for our affirmative answer is in Paper II, where specific, practical schemes for making nondemolition measurements on harmonic oscillators and free masses are discussed. All of these practical schemes are founded firmly on the fundamental principles outlined in this section. Perhaps the best thing we can do here is to indicate in a very general way the relevance of these fundamental principles to real experiments.

The objective of this section was to develop a simple, unambiguous criterion for identifying those observables of any system which, in principle, can be measured repeatedly with no uncertainty in the results. The QND condition (4.3) provides that criterion. Given this criterion, the experimenter faces a clear-cut choice. If he chooses to measure an observable other than a QND observable, he knows that, as he improves the precision of his measurements, he will eventually run "smack-dab" into an impenetrable barrier—impenetrable because it is constructed from quantum-mechanical uncertainties dictated by the uncertainty principle. On the other hand, if he chooses to measure a QND observable, he knows that nonrelativistic quantum mechanics erects no such barrier. The real value of the principles outlined in this section is that they do this job of clarifying what quantum mechanics allows.

Once the QND observables of a given system have been identified, the experimenter has a variety of options. If he is ambitious, he might try to design a measuring device which couples nearly exactly to a particular QND observable, as in continuous two-transducer measurements of X_1 (see Sec. II.F.1). This task might prove to be quite difficult, so the experimenter might rein in his ambition and choose instead to design a measuring device which couples to the QND observable only in a time-averaged sense, as in single-transducer back-action-evading measurements of X_1 (see Sec. II.F.3 and Paper II). The essential point is that all these options flow from the fundamental principles of nondemolition measurement.

Powerful, simple, clear-cut—these are words that describe the QND condition (4.3). Yet these virtues are purchased at the expense of certain assumptions about the measurement process, and under some circumstances these assumptions may make the QND condition too restrictive. Despite our belief in the utility of the QND condition, it is important to register here a couple of caveats which warn against using it carelessly.

Caveat 1. The definition of QND measurement is formulated in terms of arbitrarily precise measurements. No real experiment can achieve such perfect measurements, so the QND criterion (4.3) is always more stringent than necessary. The virtue of QND observables is that, for any desired measurement accuracy, a QND observable can do the job in principle; the caveat is that it may be possible to find an observable other than a QND observable which can also do the job.

Caveat 2. The strict operator constraint (4.3) follows from Eq. (4.2) only if one assumes that the experimenter has no control over the state of the system before the initial measurement. In most experiments this is not the case; the experimenter usually prepares the system in some way before beginning his measurements. The second caveat is that, if the experimenter does have some control over the possible initial states of the system, the QND condition (4.3) need only hold in the subspace of states which the system can actually occupy. For a simple system such as a harmonic oscillator this caveat is probably unimportant, but for more complicated systems it may make a difference.

If these caveats are kept in mind, the experimenter should be able to apply the QND condition to arbitrary systems. He can then face the experimental future free from uncertainty—about quantum-mechanical uncertainties.

APPENDIX A: CAPACITORS WITH NEGATIVE CAPACITANCE

In the text of this paper one occasionally encounters the concept of a capacitor with negative capacitance. The physical structure of such a capacitor and the details of its noise are discussed in this Appendix.

We present three models for such a capacitor. The first utilizes a mechanical spring. It will please theorists because it can be analyzed fully quantum mechanically, but it will annoy experimenters because it may not be realizable in practice. The second and third will please experimenters because they are constructed from standard electronic components; but they will annoy theorists because one (the third) functions as a negative capacitor only over a very narrow band of frequencies, and the other (the second) uses an amplifier whose internal structure is unspecified and gives a noise performance not as good as that of the first model.

In Sec. A.1 we present our first "spring-based" model capacitor; in Sec. A.2 we show that in principle it can function perfectly, introducing absolutely zero noise into the *Gedanken* experiments of Sec. III of the text; and in Sec. A.3 we present several alternative viewpoints about the nature and role of this negative capacitor. In Sec. A.4 we present our second, "amplifier-based" model capacitor; we derive an expression for the spectral density of its voltage noise; and we show that its noise is too great to do the job required in Sec. III. In Sec. A.5 we present our third, "narrow-band" negative capacitor—which also cannot do the job required in the *Gedanken* experiments of Sec. III, unless one alters them by inserting a frequency upconversion.

1. Spring-based negative capacitor

Our first "spring-based" model capacitor is shown in Fig. 7(a). It consists of three parallel plates with arbitrarily large areas. The top and bottom plates are

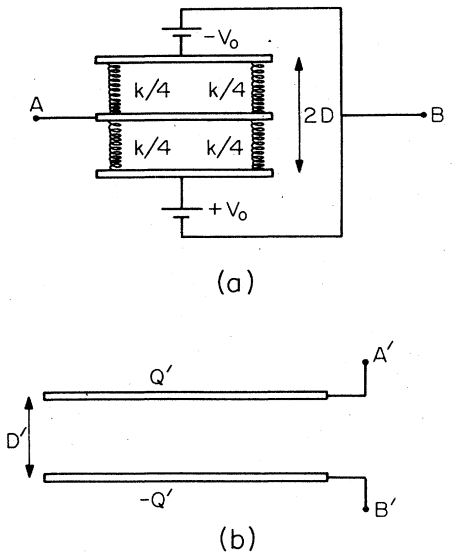

FIG. 7. (a) Model of a spring-based capacitor with negative capacitance. (b) Model for the perfect, noiseless batteries that appear in (a). For details see text.

rigidly fixed. The middle plate has negligible mass, and is free to move in response to the combined action of springs (total spring constant k) and electrostatic forces. Two batteries produce a potential difference $2V_0$ and an electric field V_0/D between the outer plates. When a charge $+Q$ moves onto the central plate from terminal A, the central plate gets pulled adiabatically upward a distance $z = V_0 Q/kD$; and terminal A thereby acquires a potential $(-V_0^2/kD^2)Q$ relative to terminal B. Thus the system functions as a capacitor with negative capacitance $-C_N$, where

$$C_N \equiv kD^2/V_0^2 . \tag{A1}$$

(The charge $Q_{\max} = \pm kD^2/V_0$, which is sufficient to drive the central plate into contact with the upper or lower plate, can be made arbitrarily large in principle while holding C_N fixed.)

This capacitor has two possible sources of noise: noise in the batteries, and noise in dynamical motions of the central plate.

The battery noise can be made zero in principle. Figure 7(b) shows a model for a noiseless dc battery. It consists of two parallel plates with finite separation D', infinitely large areas $\mathcal{A}'$ and charges $\pm Q'$, and finite surface densities of charge $\sigma' = \pm Q'/\mathcal{A}' \equiv \pm V_0/4\pi D'$. Any finite charge Q that flows through terminals A' and B' produces zero fractional change in the plate charges ($Q/Q' = 0$ since $Q' = \infty$), and therefore produces zero change in the battery voltage V_0. (Here, as throughout this paper, we ignore relativistic effects such as speed-of-light limitations on how fast the electrons can redistribute themselves on the plates near the terminals.)

Dynamical motions of the central plate of our capacitor are a delicate issue. We shall analyze them with care, first giving a heuristic semiclassical analysis and then (in Sec. A.2) giving a fully quantum-mechanical analysis. In our analysis initially we make the area $\mathcal{A}$ of the capacitor plates finite but large; the capacitance $C_0 \equiv 2(\mathcal{A}/4\pi D)$ of the central plate relative to the outer plates, finite but large; and the mass μ of the central plate, finite but tiny. The motion of the central plate is described by the dynamical variable $z(t) \equiv$ (height above central position); the charge that sits on the central plate is denoted by the dynamical variable $Q(t)$. The entire system shown in Fig. 7(a) is then described classically by the Lagrangian

$$\mathcal{L} = \tfrac{1}{2}\mu\dot{z}^2 - \tfrac{1}{2}kz^2 + \frac{V_0}{D}Qz + \frac{Q^2z^2}{2C_0D^2} - \frac{Q^2}{2C_0} . \tag{A2}$$

This Lagrangian serves two purposes: (i) its Euler-Lagrange equations $\delta\mathcal{L}/\delta z = 0$ describe the motion of the central plate; and (ii) the voltage of terminal A relative to terminal B is given by

$$V_A - V_B = -\partial\mathcal{L}/\partial Q . \tag{A3}$$

We now simplify our Lagrangian by making the plates infinitely large ($C_0 \to \infty$); we replace V_0/D by $(k/C_N)^{1/2}$ [cf. Eq. (A1)]; and we make the replacement

$$k = \mu\Omega^2 , \tag{A4}$$

where Ω is the very high natural frequency of oscillation of the central plate. The Lagrangian then reads

$$\mathcal{L} = \tfrac{1}{2}\mu\dot{z}^2 - \tfrac{1}{2}\mu\Omega^2z^2 + (\mu\Omega^2/C_N)^{1/2}Qz ; \tag{A5}$$

and the Euler-Lagrange equation of z becomes

$$\ddot{z} + \Omega^2 z = (\Omega^2/\mu C_N)^{1/2}Q . \tag{A6}$$

We assume that Ω is extremely large compared to the rate at which Q changes. Then the central plate moves nearly adiabatically in response to changes of Q:

$$z = (\mu\Omega^2 C_N)^{-1/2}Q + z_{\rm na} + z_{\rm fl} . \tag{A7}$$

Here we include a correction term $z_{\rm na}$ to account for nonadiabatic effects due to finite Ω:

$$z_{\rm na}/z \to 0 \quad \text{as} \quad \Omega \to \infty ; \tag{A8}$$

and we include a term $z_{\rm fl}$ to account semiclassically for fluctuations of the central plate demanded by quantum theory.

The voltage drop between terminals A and B, as computed from Eqs. (A3), (A5), and (A7), is

$$V_A - V_B = -(\mu\Omega^2/C_N)^{1/2}z = -Q/C_N + V_{\rm na} + V_{\rm fl} . \tag{A9a}$$

The first term is that for a perfect negative capacitor. The second, nonadiabatic term vanishes in the limit $\Omega \to \infty$:

$$V_{\rm na}/(V_A - V_B) = z_{\rm na}/z \to 0 \quad \text{as} \quad \Omega \to \infty . \tag{A9b}$$

In Sec. A.2 we shall show rigorously that, for the *Gedanken* experiments of Sec. III, the quantum fluctuations $V_{\rm fl}$ produce no charge flow in the circuit, $Q_{\rm fl} \to 0$, in the adiabatic limit $\Omega \to \infty$. The following argument explains, heuristically, why this is so: The zero-point oscillations of the central plate have a magnitude

$$|z_{\rm fl}| \sim (\hbar/\mu\Omega)^{1/2} , \tag{A10}$$

corresponding to an energy $\frac{1}{2}\hbar\Omega$. These produce a fluctuating voltage

$$|V_{\rm fl}| = (\mu\Omega^2/C_N)^{1/2}|z_{\rm fl}| \sim (\hbar\Omega/C_N)^{1/2} . \tag{A9c}$$

The characteristic frequency Ω of these fluctuations is far higher than the natural frequencies of the circuit to which our negative capacitor is hooked up. Therefore these fluctuations have great difficulty driving oscillations of the circuit:

$$Q_{\rm fl} \propto [(\text{natural frequencies})/\Omega]^2 V_{\rm fl} \propto \Omega^{-3/2} \to 0 \text{ as } \Omega \to \infty . \tag{A11}$$

In summary, our heuristic argument shows that the model negative capacitor of Fig. 7 functions perfectly (no noise) in the *Gedanken* experiments of Sec. III. However, this is so only in the idealized limits that (i) the area of the capacitor's plates is infinite ($\mathcal{A} \to \infty$, $C_0 \to \infty$); (ii) the natural oscillation frequency of its central plate is infinite ($\Omega^2 = k/\mu \to \infty$); and (iii) one ignores relativistic corrections, issues of strengths of materials, etc.

2. *Gedanken* experiment to measure the momentum of a free mass

We now sketch a fully quantum-mechanical analysis of one of the *Gedanken* experiments of Sec. III, replacing the ideal negative capacitance of Sec. III by the spring-based model negative capacitance of Fig. 7(a). The *Gedanken* experiment we choose is the measurement of the momentum of a free mass (Sec. III.A.2). The reader can perform a similar calculation for the *Gedanken* experiment to measure the X_1 of an oscillator (Sec. III.B). The result will be the same: In the adiabatic limit $\Omega \to \infty$, the negative capacitor produces zero noise.

The physical setup of our momentum-measuring experiment is that of Fig. 6(a) with (i) the noisy amplifier (dashed part) removed; (ii) the capacitance C set to infinity; and (iii) our negative capacitor (Fig. 7) inserted at the location of the dotted arrow. The Lagrangian of everything except the negative capacitor is Eq. (3.3); the contribution of the negative capacitor is Eq. (A5); and the value of the negative capacitance which we require to convert our velocity sensor into a momentum sensor is

$$-C_N = -1/mK^2 . \tag{A12}$$

[See the sentence preceding Eq. (3.9).] The total Lagrangian then becomes

$$\mathcal{L} = \tfrac{1}{2} m\dot{x}^2 + \tfrac{1}{2} L\dot{Q}^2 - Km\dot{x}Q + Fx + \tfrac{1}{2}\mu\dot{z}^2 - \tfrac{1}{2}\mu\Omega^2 z^2 + (m\mu K^2\Omega^2)^{1/2} Qz . \tag{A13}$$

We shall see that, in the limit $\Omega \to \infty$, this Lagrangian gives the same quantum-mechanical results for the measurement of the momentum $\hat{p}$, and force F, as did the Lagrangian (3.9) which contained a perfect negative capacitor $-C_N = -1/mK^2$.

The canonical momenta for the Lagrangian (A13) are

$$p = \frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x} - KmQ , \quad \Pi = \frac{\partial \mathcal{L}}{\partial \dot{Q}} = L\dot{Q} , \quad \pi = \frac{\partial \mathcal{L}}{\partial \dot{z}} = \mu\dot{z} . \tag{A14}$$

The Hamiltonian $H = p\dot{x} + \Pi\dot{Q} + \pi\dot{z} - \mathcal{L}$, after quantization, is

$$\hat{H} = \frac{\hat{p}^2}{2m} + \left(\frac{\hat{\Pi}^2}{2L} + \tfrac{1}{2} mK^2\hat{Q}^2\right) + \left(\frac{\hat{\pi}^2}{2\mu} + \tfrac{1}{2}\mu\Omega^2\hat{z}^2\right) - F\hat{x} + K\hat{p}\hat{Q} - (m\mu K^2\Omega^2)^{1/2}\hat{Q}\hat{z} . \tag{A15}$$

This Hamiltonian will give the same results, when $\Omega \to \infty$, as did the Hamiltonian (3.10) with a perfect negative capacitor.

The Heisenberg equations for the Hamiltonian (A15) are

$$\frac{d\hat{p}}{dt} = F , \quad \frac{d\hat{x}}{dt} = \frac{\hat{p}}{m} + K\hat{Q} ,$$
$$\frac{d\hat{\pi}}{dt} = -\mu\Omega^2\hat{z} + (m\mu K^2\Omega^2)^{1/2}\hat{Q} , \quad \frac{d\hat{z}}{dt} = \frac{\hat{\pi}}{\mu} , \tag{A16}$$
$$\frac{d\hat{\Pi}}{dt} = -K\hat{p} - mK^2\hat{Q} + (m\mu K^2\Omega^2)^{1/2}\hat{z} , \quad \frac{d\hat{Q}}{dt} = \frac{\hat{\Pi}}{L} .$$

These equations describe coupled, driven harmonic oscillators. They can be decoupled by the change of variables

$$\hat{y}_1 = \hat{Q} + \frac{(m\mu)^{1/2}K}{L\Omega}\hat{z} , \quad \hat{y}_2 = \hat{z} - \frac{(m/\mu)^{1/2}K}{\Omega}\hat{Q} . \tag{A17}$$

Here $\hat{y}_1$ has eigenfrequency zero, and in the adiabatic limit ($\Omega \to \infty$) it becomes $\hat{Q}$; $\hat{y}_2$ has eigenfrequency

$$\sigma \equiv (\Omega^2 + mK^2/L)^{1/2} , \tag{A18}$$

and in the adiabatic limit it becomes $\hat{z}$. By changing variables to $\hat{y}_1$, $\hat{y}_2$, then solving the Heisenberg equations, and then rewriting the solution in terms of $\hat{Q}$ and $\hat{z}$ we obtain

$$\hat{p}(t) = \hat{p}_0 + Ft , \tag{A19a}$$

$$\hat{Q}(t) = \left(\frac{\Omega}{\sigma}\right)^2 \left\{ \left(\hat{Q}_0 + \frac{(m\mu)^{1/2}K}{L\Omega}\hat{z}_0\right) + \left(\frac{\hat{\Pi}_0}{L} + \frac{(m\mu)^{1/2}K}{L\Omega}\frac{\hat{\pi}_0}{\mu}\right)t + \left(\frac{mK^2}{L\Omega^2}\hat{Q}_0 - \frac{(m\mu)^{1/2}K}{L\Omega}\hat{z}_0\right)\cos\sigma t \right.$$
$$\left. + \left(\frac{mK^2}{L\Omega^2}\frac{\hat{\Pi}_0}{L} - \frac{(m\mu)^{1/2}K}{L\Omega}\frac{\hat{\pi}_0}{\mu}\right)\frac{\sin\sigma t}{\sigma} - \frac{K}{L}\left(\tfrac{1}{2}\hat{p}_0 t^2 + \tfrac{1}{6}Ft^3\right) - \frac{K}{L}\frac{mK^2}{L\Omega^2}\left[\frac{\hat{p}_0}{\sigma^2}(1-\cos\sigma t) + \frac{F}{\sigma^2}\left(t - \frac{\sin\sigma t}{\sigma}\right)\right]\right\} , \tag{A19b}$$

$$\hat{z}(t) = \left(\frac{\Omega}{\sigma}\right)^2 \left\{ \left(\hat{z}_0 - \frac{(m/\mu)^{1/2}K}{\Omega}\hat{Q}_0\right)\cos\sigma t + \left(\frac{\hat{\pi}_0}{\mu} - \frac{(m/\mu)^{1/2}K}{\Omega}\frac{\hat{\Pi}_0}{L}\right)\frac{\sin\sigma t}{\sigma} + \left(\frac{mK^2}{L\Omega^2}\hat{z}_0 + \frac{(m/\mu)^{1/2}K}{\Omega}\hat{Q}_0\right) \right.$$
$$\left. + \left(\frac{mK^2}{L\Omega^2}\frac{\hat{\pi}_0}{\mu} + \frac{(m/\mu)^{1/2}K}{\Omega}\frac{\hat{\Pi}_0}{L}\right)t + \frac{(m/\mu)^{1/2}K^2}{L\Omega}\left[\frac{\hat{p}_0}{\sigma^2}(1-\cos\sigma t) + \frac{F}{\sigma^2}\left(t - \frac{\sin\sigma t}{\sigma}\right) - \tfrac{1}{2}\hat{p}_0 t^2 - \tfrac{1}{6}Ft^3\right]\right\} . \tag{A19c}$$

The remaining variables can easily be computed from these using the Heisenberg equations (A16). In these solutions a subscript "0" means the value at $t=0$:

$$\hat{p}_0 \equiv \hat{p}(0), \quad \hat{Q}_0 \equiv \hat{Q}(0), \text{ etc.} \tag{A20}$$

The solution (A19b) for $\hat{Q}(t)$ illustrates, fully quantum mechanically, the phenomena sketched semiclassically in the preceding section: (i) In the exact adiabatic limit $\Omega \to \infty$, the charge $\hat{Q}(t)$ that flows in the circuit is identical to that obtained with a perfect, noiseless negative capacitor [compare Eqs. (A19b) and (3.12)]. (ii) When Ω is finite but $\Omega \gg (mK^2/L)^{1/2}$ = (natural frequency of circuit without negative capacitor), there are corrections in $\hat{Q}(t)$ due to nonadiabatic behavior; but these corrections vanish as $\Omega \to \infty$.

Quantum-mechanical fluctuations in $\hat{Q}(t)$ show up when one computes the variance $\Delta Q(t) = \langle (\hat{Q} - \langle \hat{Q} \rangle)^2 \rangle^{1/2}$ in terms of the variances at time $t=0$. Because $\hat{Q}(t)$ reduces to the "perfect-capacitor" form [Eq. (3.12)] when $\Omega \to \infty$, we are guaranteed that $\Delta Q(t)$ will reduce to the perfect-capacitor variance [Eq. (3.14)] when $\Omega \to \infty$. Thus, in the adiabatic limit, quantum fluctuations of the central plate have no effect on the charge $\langle \hat{Q}(t) \rangle$ that flows, or on its variance $\Delta Q(t)$. Our negative capacitor does its job perfectly and noiselessly.

3. Alternative viewpoints on the spring-based negative capacitor

We have argued in the text (Sec. III.A.2) that, in monitoring the motion of a mechanical system, a momentum sensor is equivalent to a velocity sensor plus a negative capacitor. Similarly (Appendix B.2), in monitoring an electromagnetic system, a sensor for generalized momentum is equivalent to a sensor for generalized velocity plus a negative spring.

In designing practical variants of such sensors, it may be useful to keep in mind several different viewpoints about negative capacitors and negative springs. One viewpoint is that embodied in the phrases "capacitor with negative capacitance" and "spring with negative spring constant." Two other viewpoints are presented in this section.

Our second viewpoint on negative capacitors is this (the extension to negative springs should be obvious): A velocity sensor is equivalent to a momentum sensor plus a restoring force in the sensor's circuit [term $\frac{1}{2}mK^2\hat{Q}^2$ in the Hamiltonian of Eq. (3.5)]. The stronger is the coupling of the velocity sensor to the mechanical mass (coupling constant K), the stronger is the restoring force in the sensor's circuit. If one wishes to measure the mechanical momentum more accurately than the standard quantum limit, one must make K so strong that the restoring force causes the circuit to oscillate through several cycles during the measurement time τ. Because of these oscillations, the effects of the driving signal (voltage $-Kp$) do not accumulate monotonically in the circuit. Consequently, the signal-to-noise ratio is debilitated, and the measurement cannot beat the standard quantum limit [cf. Eq. (3.8)]. To rectify the situation one must modify the sensing circuit so that it contains a low-frequency ($f \lesssim 1/\tau$) normal mode in which the signal can accumulate monotonically. Our so-called "spring-based negative capacitor" accomplishes just this. It gives the readout circuit two dynamical degrees of freedom instead of one; and when it is properly tuned to the rest of the sensor [$kD^2/V_0^2 = 1/mK^2$; Eqs. (A1) and (A12)], one of the degrees of freedom $\{\hat{y}_1 = \hat{Q} + [(m\mu)^{1/2}K/L\Omega]\hat{z}\}$ has zero eigenfrequency. The signal builds up monotonically in this degree of freedom giving, in principle, an arbitrarily large signal-to-noise ratio.

Our third viewpoint on negative capacitors builds on this second viewpoint. When our "spring-based negative capacitor" is included in the sensor, then the sensor circuit has two normal modes. It is essential that one of the normal modes have a low enough eigenfrequency, $f \lesssim 1/\tau$, for the signal to accumulate monotonically. However, it is *not* essential that the other normal mode have such a high eigenfrequency that it decouples from the rest of the system ($\Omega \to \infty$; adiabatic limit; situation assumed in all previous discussion). For example, we might let Ω, the natural frequency of the central plate in the "negative capacitor," be of order $(mK^2/L)^{1/2}$, the natural frequency of the circuit in the absence of the negative capacitor:

$$\Omega \simeq (mK^2/L)^{1/2}. \tag{A21}$$

Then, it turns out, the mechanical motion of the central plate $\hat{z}(t)$ is influenced sufficiently by the zero-frequency normal mode $\hat{y}_1(t)$ that one can read out from that motion the signal contained in $y_1(t)$. More specifically, for the *Gedanken* experiment of Sec. III.A.2, with initial conditions

$$\begin{aligned} &\langle \hat{\Pi}_0 \rangle = \langle \hat{Q}_0 \rangle = \langle \hat{\pi}_0 \rangle = \langle \hat{z}_0 \rangle = 0, \quad \langle \hat{p}_0 \rangle = p_0, \\ &\Delta\Pi_0 = (\hbar L/2\tau)^{1/2}, \quad \Delta Q_0 = (\hbar\tau/2L)^{1/2}, \\ &\Delta\pi_0 = (\hbar\mu/2\tau)^{1/2}, \quad \Delta z_0 = (\hbar\tau/2\mu)^{1/2}, \quad \Delta p_0 = 0, \\ &\text{no correlations of above variables,} \end{aligned} \tag{A22}$$

the expectation value and variance of the central plate's position at time τ are [Eq. (A19c)]

$$\langle \hat{z}(\tau) \rangle = -\frac{1}{2}\left(\frac{\Omega}{\sigma}\right)^2 \frac{(m/\mu)^{1/2}K^2}{L\Omega} p_0 \tau^2 \left[1 + O\left(\frac{1}{\Omega^2\tau^2}\right)\right], \tag{A23a}$$

$$\Delta z(\tau) \simeq (\Omega/\sigma)^2 (\hbar\tau/\mu)^{1/2}[1 + O(1/\Omega^2\tau^2)]. \tag{A23b}$$

Here use has been made of Eq. (A21), and for simplicity the classical driving force has been omitted ($F=0$). One can attach a pointer with a scale to the central plate, and in principle one can read out $z(\tau)$ from that pointer with probable error $\Delta z(\tau)$. From the result one can infer the free-mass momentum p_0 to within probable error

$$\delta p_0 = \frac{\Delta z(\tau)}{\partial \langle \hat{z}(\tau) \rangle / \partial p_0} \simeq \left(\frac{\hbar L}{K^2\tau^3}\right)^{1/2} \cdot \left[1 + O\left(\frac{L}{mK^2\tau^2}\right)\right], \tag{A24}$$

where again Eq. (A21) has been used. For a given τ, if the coupling is stronger than $K^2 = L/m\tau^2$, then the measurement can be more accurate than the standard quantum limit [$\delta p_0 < (\hbar m/\tau)^{1/2}$]; and if $K \to \infty$, then the measurement can be arbitrarily accurate.

In this variant of the experiment the "spring-based negative capacitor" functions as a readout device ("charge meter") which is carefully tuned [$kD^2/V_0^2 = 1/mK^2$; Eqs. (A1) and (A12)] to the rest of the sensing

circuit. The pointer attached to the central plate gets displaced by an amount $\langle\hat{z}(\tau)\rangle$, which is proportional to the charge that has flowed onto the central plate—and thence proportional to the momentum p_0 of the free mass. A person adopting this "third viewpoint" should realize that the careful tuning ($kD^2/V_0^2 = 1/mK^2$) is required to produce a zero-frequency mode in which the signal can accumulate ("viewpoint two"), but he need not be aware that his charge meter is functioning, in effect, like a negative capacitor ("viewpoint one").[11]

4. Amplifier-based negative capacitor

Figure 8(a) shows a model negative capacitor constructed from standard electronic components, including a voltage amplifier with infinite input impedance. The amplifier has arbitrarily large amplification at all frequencies of interest, and its equivalent voltage and current noise sources $V_n(t)$ and $I_n(t)$ have spectral densities constrained by the quantum limit

$$S_V(f)S_I(f) \gtrsim (4\pi\hbar f)^2 \tag{A25}$$

[Heffner (1962); Eq. (3.7) of this paper]. (For simplicity we assume zero correlation between the voltage and current noises.) The capacitors C_1 and C_2 act as a voltage divider. When a voltage V is applied to A, the amplifier forces a negative charge $-C_0(C_1/C_2)V$ to flow in at terminal A and onto the capacitor C_0. Thus the system exhibits a negative capacitance $-C_N$ given by

$$C_N = C_0(C_1/C_2)\,. \tag{A26}$$

It is straightforward to show that the voltage-current relation for this device is

$$\tilde{V} = \frac{\tilde{I}}{i2\pi f C_N} - \tilde{V}_n\left(1+\frac{C_2}{C_1}\right) + \frac{\tilde{I}_n}{i2\pi f}\left(\frac{1}{C_N}+\frac{1}{C_1}\right), \tag{A27}$$

where a tilde denotes Fourier transform at frequency $f[\tilde{A}(f) = \int_{-\infty}^{\infty} A(t)e^{i2\pi ft}\,dt]$. This is identical to the voltage-current relation for the "Thevenin equivalent circuit" of Fig. 7(b). The voltage noise source $V_N(t)$ for that circuit has spectral density which we can read off Eq. (A27):

$$S_{V_N}(f) = \left(1+\frac{C_2}{C_1}\right)^2 S_V(f) + \frac{S_I(f)}{(2\pi f)^2}\left(\frac{1}{C_N}+\frac{1}{C_1}\right)^2. \tag{A28}$$

This noise is minimized for fixed C_N by setting $C_1/C_2 \to \infty$, $C_1/C_N \to \infty$, and by impedance matching the amplifier so that $S_I/S_V = (2\pi f C_N)^2$. (In principle the impedance matching can be achieved at any chosen frequency by a transformer that immediately precedes the amplifier input). Then the spectral density of the equivalent noise source V_N becomes

$$S_{V_N} = (S_V S_I)^{1/2}/\pi f C_N\,. \tag{A29}$$

The quantum limit (A25) for the amplifier then implies

$$S_{V_N} \gtrsim 4\hbar/C_N\,. \tag{A30}$$

This is the very best noise performance that the model negative capacitor of Fig. 8 can possibly achieve. It is instructive to compare this noise, which has a white spectrum, with that of our spring-based model for a negative capacitor [Eq. (A9c)], which is concentrated at the very high frequency Ω.

Unfortunately, the noise performance (A30) is too poor to permit use of this negative capacitor in the "arbitrarily quick and accurate" *Gedanken* experiments of Sec. III. For example, in the momentum measuring experiment of Sec. III.A.2 we require $C_N = 1/mK^2$, where m is the mass of the "free mass" being measured, and K is the coupling constant in the transducer. In dc measurements of duration τ our model capacitor would superimpose on the transducer output a fluctuating voltage with variance

$$\Delta V_N \simeq \left(S_{V_N}\frac{1}{2\tau}\right)^{1/2} \gtrsim \left(\frac{2\hbar}{C_N\tau}\right)^{1/2} = K\left(\frac{2\hbar m}{\tau}\right)^{1/2}. \tag{A31}$$

For comparison, the signal voltage produced by the transducer is $V_s = -Kp$ [cf. Eq. (3.11) with $V_s = d\Pi/dt$], where p is the momentum to be measured. Evidently the voltage noise V_N of the negative capacitor produces an uncertainty

$$\delta p \simeq (2\hbar m/\tau)^{1/2} \tag{A32a}$$

in one's measurement of p, and a corresponding uncertainty

$$\delta F \simeq \delta p/\tau \simeq (2\hbar m/\tau^3)^{1/2} \tag{A32b}$$

in one's knowledge of any classical force acting on the

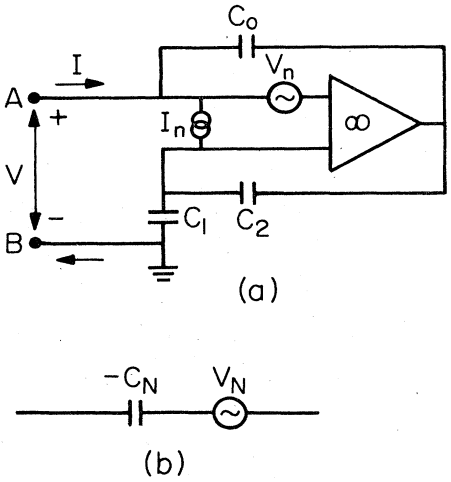

FIG. 8. (a) Model of an amplifier-based capacitor with negative capacitance. (b) Thevenin equivalent circuit for this model negative capacitor. For details see text.

[11] In our original article in *Physical Review Letters* (Thorne *et al.*, 1978), we discussed a *Gedanken* experiment for an arbitrarily quick and accurate back-action-evading measurement of the X_1 of an electromagnetic oscillator (cf. Appendix B.2 of this paper). In that discussion we asserted that a torque Γ in the sensing system could be read out with precision $\Delta\Gamma \simeq (I\hbar/\tau^3)^{1/2}$. We had invented the required "torque-balance readout system" at the time of our Letter (though we did not describe it in the Letter). Our viewpoint on that torque balance was the third viewpoint described above; and we were unaware that our balance was functioning, in effect, like a negative spring. The reason we presented in Thorne *et al.* (1978) a *Gedanken* experiment for measuring an electromagnetic oscillator, rather than a mechanical oscillator, was that we had not yet invented the "negative-capacitor" readout system of Fig. 7.

free mass. These uncertainties are equal to the standard quantum limit for a free mass. Thus the noise of our second negative capacitor [Fig. 8(a)] is too great to permit its use in measurements designed to beat the standard quantum limit.

5. Narrow-band negative capacitor

When one is performing measurements in a narrow band of angular frequencies $\Delta\omega$ around a high "carrier" frequency Ω, one can use an inductor as a negative capacitor. Aside from fractional corrections of order $\Delta\omega/\Omega$, an inductor with inductance $L \equiv (\Omega^2 C_N)^{-1}$ has the same impedance in this band as a negative capacitor $-C_N$:

$$Z = -i2\pi f L = [-i2\pi f(-C_N)]^{-1}[1 + O(\Delta\omega/\Omega)]. \qquad \text{(A33)}$$

In principle such a "narrow-band" negative capacitor can be noiseless.

The "arbitrarily quick and accurate" *Gedanken* experiments of Sec. III require a negative capacitor that operates over a broad band of frequencies, $0 < f \lesssim 1/2\tau$. Thus an inductor cannot do the required job. However, one can invent a more complicated version of those *Gedanken* experiments, in which, for a measurement of the momentum of a free mass, the output of the velocity transducer is multiplied by $\cos\Omega t$ with $\Omega \gg 1/\tau$. Similarly, for a measurement of the X_1 of an oscillator, the outputs of both the coordinate and velocity transducers can be multiplied by $\cos\Omega t$. Then the readout is at frequencies $\simeq \Omega$ in the band $\Delta\omega \simeq \pi/\tau$, and a narrow-band negative capacitor (i.e., an inductor) does an adequate job of converting the velocity transducer into a momentum transducer. Such a measurement can determine the momentum of a free mass with accuracy $\delta p_0 \simeq (\Omega\tau)^{-1/2}(\hbar m/\tau)^{1/2}$, or the X_1 of an oscillator with accuracy $\delta X_1 \simeq (\omega\tau)^{-1/2}(\Omega\tau)^{-1/2}(\hbar/2m\omega)^{1/2}$. This trick of "upconversion" of the signal to a carrier frequency Ω is discussed in detail in Paper II.

APPENDIX B: PHYSICAL REALIZATIONS OF HAMILTONIAN (3.16) FOR ARBITRARILY QUICK AND ACCURATE MEASUREMENTS OF X_1

Here we describe *Gedanken* apparatus by which, in principle, one could make the "arbitrarily quick and accurate" measurements of X_1 described abstractly in Sec. III.B. Our objective is *not* to describe practical apparatus for real experiments. (Practical issues are discussed in Paper II.) Rather, we seek to demonstrate, in the manner of a mathematician proving a theorem, that in principle there can exist apparatus governed precisely by the Hamiltonian of Sec. III.B [Eq. (3.16)].

Section B.1 of this Appendix describes apparatus for measuring a mechanical oscillator, and discusses the relationship between classical generators to be used in the apparatus and quantum mechanical generators. Section B.2 describes apparatus for an electromagnetic oscillator.

1. Mechanical oscillator

a. Physical description

Figure 9 shows a physical realization of the coupled oscillator and measuring apparatus which were de-

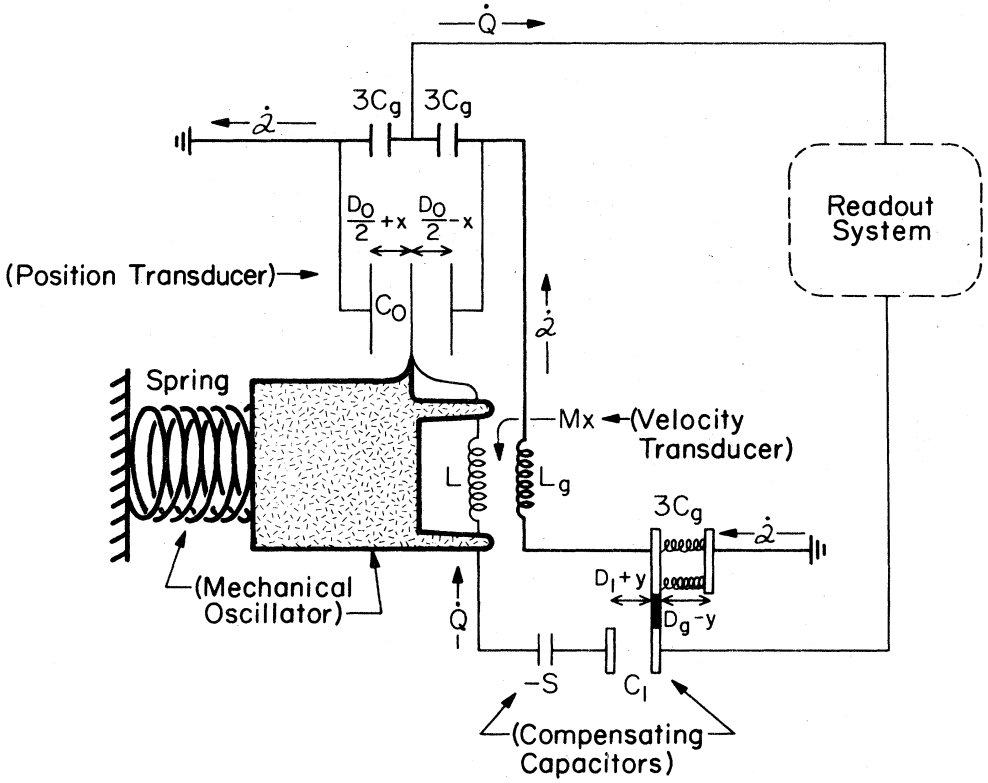

FIG. 9. Idealized physical realization of a system for measuring the $\hat{X}_1$ of a mechanical oscillator arbitrarily quickly and accurately. See text for discussion.

scribed abstractly in Sec. III.B. In this figure our *mechanical oscillator* is drawn with very thick lines. It consists of a mass (stippled square) coupled by a spring to a rigid wall.

Our *electromagnetic generator* circuit is drawn with lines of medium thickness. It is an LC circuit with a single lumped inductance L_g and with total capacitance C_g split up among three capacitors in series—two at the top of the diagram; the third at the lower right-hand side. This generator will be excited into a highly classical, coherent state, thereby producing voltages proportional to $\cos\omega t$ across its capacitors, and a current proportional to $-\sin\omega t$ through its inductor. These voltages and this current will provide the sinusoidal couplings for our position and momentum transducers.

The *meter* of Sec. III.B (a circuit with self-inductance L but no net capacitance) and the *transducers* which couple the meter to the mechanical oscillator are drawn with thin lines. The *position tranducer* is the three-plate balanced capacitor labeled C_0 in the upper part of Fig. 9. The outer plates will be biased with voltages $\pm$ (const) $\cos\omega t$ by the generator's capacitors $3C_g$; and as a result the central plate, which is attached rigidly to the oscillator, will acquire a voltage proportional to $x\cos\omega t$. The *momentum transducer* consists of two parts: a *velocity transducer* [mutual inductance Mx between L and L_g which, because of the generator current $I_g \propto \sin\omega t$ through L_g, will produce a voltage across L that is proportional to $d(x\sin\omega t)/dt = \dot{x}\sin\omega t + \omega x\cos\omega t$]; and a complicated system of *compensating capacitors* which convert the velocity transducer into a momentum transducer [net output voltage proportional to $(p/m\omega)\sin\omega t + x\cos\omega t$]. We adjust the relative strengths of the couplings in our position and momentum transducers so that the total signal voltage in the meter (thin-line circuit of Fig. 9) is $K\cdot[x\cos\omega t - (p/m\omega)\sin\omega t] \equiv K\cdot X_1$.

The *readout system* measures the total charge Q that the signal voltage KX_1 has driven through the meter circuit. In the limit of very strong coupling, we can put the quantum-classical cut between the meter and the readout system, and we can forego any detailed mathematical description of the readout system; cf. Sec. III.B and Appendix C.2.

b. Derivation of the Hamiltonian

Initially we analyze the system of Fig. 9 in the Lagrangian formalism of classical mechanics; then we compute the Hamiltonian and quantize it.

In the Lagrangian formalism the *mechanical oscillator* is characterized by its mass m, frequency ω, and time-dependent position $x(t)$. The *electromagnetic generator*, which produces the sinusoidal couplings, is characterized by its total capacitance C_g and inductance L_g, and by the current $\dot{\mathcal{Q}}(t) \equiv d\mathcal{Q}/dt$ which flows in it. The eigenfrequency of the generator is identical to that of the mechanical oscillator:

$$L_g C_g = 1/\omega^2 . \tag{B1}$$

The *meter* is characterized by its self-inductance L and its current $\dot{Q}(t) = dQ/dt$.

From the constant K, which characterizes the coupling of the oscillator to the meter, and from the mass m and frequency ω of the oscillator, we can construct a characteristic length scale s:

$$s \equiv m\omega^2/K^2 . \tag{B2a}$$

We shall choose the generator's capacitance C_g to be huge compared with s, and we shall introduce the small dimensionless parameter

$$\varepsilon \equiv s/3C_g . \tag{B2b}$$

Before each measurement the generator, regarded quantum mechanically, will be excited into a coherent state with

$$\langle\hat{\mathcal{Q}}\rangle = \mathcal{Q}_0\cos\omega t , \quad \Delta\mathcal{Q} = (\hbar/2L_g\omega)^{1/2} , \tag{B3a}$$

$$\mathcal{Q}_0 \equiv Ks^2/\varepsilon^{3/2} . \tag{B3b}$$

The mean number of quanta in the generator, $\langle\hat{N}\rangle \equiv N_0$, and the fractional width of its wave packet will then be

$$\langle\hat{N}\rangle = N_0 = \frac{\mathcal{Q}_0^2/2C_g}{\hbar\omega} = \frac{3}{2}\frac{m(\omega s)^2}{\varepsilon^2\hbar\omega} , \tag{B3c}$$

$$\Delta\mathcal{Q}/\mathcal{Q}_0 = \tfrac{1}{2}N_0^{-1/2} . \tag{B3d}$$

In the limit $\varepsilon \to 0$, the generator will contain an infinite number of quanta, $N_0 \to \infty$, and it will become fully classical, $\Delta\mathcal{Q}/\mathcal{Q}_0 \to 0$. We shall keep ε finite but small in our analysis, until we have obtained our Hamiltonian. Then (Sec. B.1.c) we shall take $\varepsilon \to 0$, thereby bringing our Hamiltonian into the form [Eq. (3.16)] studied in Sec. III.B and Appendix C.

We now construct the Lagrangian for our system, choosing the magnitudes of various parameters along the way so that in the limit $\varepsilon \to 0$ the corresponding Hamiltonian will reduce to Eq. (3.16).

The mechanical oscillator (thick lines in Fig. 9) has the familiar Lagrangian

$$\mathcal{L}_0 = \tfrac{1}{2}m\dot{x}^2 - \tfrac{1}{2}m\omega^2x^2 . \tag{B4}$$

The generator's inductance L_g is fixed in inertial space. The meter's inductance L is partly attached to the mechanical oscillator, and partly attached to inertial space—with the details of the attachments designed to produce a mutual inductance between L and L_g which is proportional to the oscillator's displacement x. The proportionality constant M is chosen to be

$$M = \varepsilon^{3/2}(\omega s)^{-2} . \tag{B5}$$

The resulting Lagrangian associated with the inductances is

$$\mathcal{L}_i = \tfrac{1}{2}L\dot{Q}^2 + \tfrac{1}{2}L_g\dot{\mathcal{Q}}^2 + Mx\dot{Q}\dot{\mathcal{Q}} \tag{B6a}$$

$$= \tfrac{1}{2}L\dot{Q}^2 + \tfrac{1}{2}L_g\dot{\mathcal{Q}}^2 + Kx(\dot{Q}/\omega)(\dot{\mathcal{Q}}/\omega\mathcal{Q}_0) . \tag{B6b}$$

Consider next the circuitry above the mechanical oscillator in Fig. 9—i.e., the position transducer, plus two-thirds of the generator's capacitance. The two capacitors labeled $3C_g$ are fixed in inertial space, as are the outside two plates of the capacitor C_0. The central plate of C_0 is rigidly attached to the mechanical oscillator, so that its separations from the outside plates are $\frac{1}{2}D_0 \pm x$. We define C_0 to be the capacitance between the outside plates at a moment when there is no charge on the central plate ($Q = 0$). We set

$$C_0 \equiv s/\varepsilon^{3/4}, \quad D_0 \equiv s/\varepsilon^{1/2}, \tag{B7}$$

so that in the limit $\varepsilon \to 0$, (i) the plate separation D_0 gets arbitrarily large, leaving plenty of room for the oscillator to move, and (ii) the linear size of the plates, $(D_0 C_0)^{1/2} \propto \varepsilon^{-5/8}$, gets far larger than their separation, $D_0 \propto \varepsilon^{-1/2}$. The total energy in the capacitors, expressed in terms of Q, $\mathcal{Q}$, and x, is equal to minus the Lagrangian, $-\mathcal{L}_c$, of the capacitors. A straightforward computation gives

$$\mathcal{L}_c = -\frac{\mathcal{Q}^2}{3C_g(1+2C_0/3C_g)} - \frac{Q^2}{8C_0}\left(1+\frac{2C_0}{3C_g}-\frac{4(x/D_0)^2}{1+2C_0/3C_g}\right) - \frac{2\mathcal{Q}Qx/D_0}{3C_g(1+2C_0/3C_g)}. \tag{B8a}$$

By using Eqs. (B7), (B2b), and (B3b), and discarding all contributions to $\mathcal{L}_c$ which vanish in the limit $\varepsilon \to 0$, we bring this into the form

$$\mathcal{L}_c = -\mathcal{Q}^2/3C_g(1+2\varepsilon^{1/4}) - 2KxQ(\mathcal{Q}/\mathcal{Q}_0). \tag{B8b}$$

[Eqs. (B2) and (B3) imply $\mathcal{Q}_0^2/C_g \propto 1/\varepsilon^2$, which forces us to keep $\varepsilon^{1/4}$ correction in Eq. (B8b).]

The mutual inductance of Eq. (B6) produces a time-dependent velocity coupling. As in Sec. III.A.2, so also here, a negative capacitor is needed to convert this velocity coupling into a pure momentum coupling—but now the negative capacitance must be time dependent. It is achieved by the compensating capacitors at the bottom of Fig. 9. These include (i) a constant negative capacitance $-s$, which has the internal structure discussed in Appendix A.1 and which contributes

$$\mathcal{L}_s = +\tfrac{1}{2}Q^2/s \tag{B9}$$

to the Lagrangian; and (ii) the variable positive capacitor "C_1." The left plate of C_1 is fixed in inertial space and the right plate of C_1 is attached rigidly by an insulator to the movable left plate of the generator's capacitor "$3C_g$"—which in turn is attached by insulated springs (total spring constant k) to the right plate of $3C_g$. This arrangement enables the generator to modulate the plate separation of C_1 and thereby modulate its capacitance. The total mass of the movable plates is vanishingly small (eigenfrequency of vibration infinitely large) so that, like the central plate of the model negative capacitor in Fig. 7(a), they move adiabatically and they inject zero noise into the electrical system. When no charge is on the capacitors, the movable plates have position $y=0$ and the capacitances are $3C_g$ and C_1. When charges $\mathcal{Q}$ and Q are applied, the equilibrium position is

$$y = (1/k)(\mathcal{Q}^2/6C_g D_g - Q^2/2C_1 D_1). \tag{B10}$$

We set

$$k = \tfrac{1}{2}\varepsilon^{-3/2} m\omega^2, \quad D_g = \varepsilon^{-5/8} s, \quad C_1 = \varepsilon^{-1/8} s, \quad D_1 = \varepsilon^{1/4} s, \tag{B11}$$

so that in the limit $\varepsilon \to 0$ the linear sizes of the plates become large compared to their equilibrium separations and large compared to their displaced separations:

$$D_g/(C_g D_g)^{1/2} \to 0, \quad (D_g - y)/(C_g D_g)^{1/2} \to 0,$$
$$D_1/(C_1 D_1)^{1/2} \to 0, \quad (D_1 + y)/(C_1 D_1)^{1/2} \to 0.$$

A straightforward computation gives for the Lagrangian of these two variable capacitors (equal to minus the energy in the springs and capacitors)

$$\mathcal{L}_{vc} = -\frac{\mathcal{Q}^2}{6C_g}\left(1-\frac{1}{12}\frac{\mathcal{Q}^2}{kC_g D_g^2}\right) - \frac{Q^2}{2C_1}\left(1-\frac{1}{4}\frac{Q^2}{kC_1 D_1^2}\right) - \frac{1}{12}\frac{\mathcal{Q}^2 Q^2}{kC_1 D_1 C_g D_g}. \tag{B12a}$$

Using Eqs. (B11), (B2), and (B3b), and discarding terms in $\mathcal{L}_{vc}$ that vanish when $\varepsilon \to 0$, we bring this into the form

$$\mathcal{L}_{vc} = -\frac{\mathcal{Q}^2}{6C_g}\left[1-\tfrac{1}{2}\varepsilon^{3/4}\left(\frac{\mathcal{Q}}{\mathcal{Q}_0}\right)^2\right] - \frac{Q^2}{2s}\left(\frac{\mathcal{Q}}{\mathcal{Q}_0}\right)^2. \tag{B12b}$$

(The $\varepsilon^{3/4}$ correction must be kept here because $\mathcal{Q}_0^2/C_g \propto 1/\varepsilon^2$.)

The total Lagrangian is the sum of Eqs. (B4), (B6b), (B8b), (B9), and (B12b):

$$\mathcal{L} = \tfrac{1}{2}m\dot{x}^2 - \tfrac{1}{2}m\omega^2 x^2 + \tfrac{1}{2}L_g\dot{\mathcal{Q}}^2 - \frac{\mathcal{Q}^2}{2C_g}\left[\frac{1+\frac{2}{3}\varepsilon^{1/4}}{1+2\varepsilon^{1/4}} - \tfrac{1}{6}\varepsilon^{3/4}\left(\frac{\mathcal{Q}}{\mathcal{Q}_0}\right)^2\right] + \tfrac{1}{2}L\dot{Q}^2 + \frac{Q^2}{2s}\left[1-\left(\frac{\mathcal{Q}}{\mathcal{Q}_0}\right)^2\right] - 2KxQ\left(\frac{\mathcal{Q}}{\mathcal{Q}_0}\right) + Kx\left(\frac{\dot{Q}}{\omega}\right)\left(\frac{\dot{\mathcal{Q}}}{\omega\mathcal{Q}_0}\right). \tag{B13}$$

The terms multiplying $-\mathcal{Q}^2/2C_g$ produce only a slight renormalization of the generator frequency and a slight anharmonicity in its oscillations—and these effects vanish in the limit $\varepsilon \to 0$. Therefore we may discard these terms, thereby bringing our Lagrangian into the form

$$\mathcal{L} = \tfrac{1}{2}m\dot{x}^2 - \tfrac{1}{2}m\omega^2 x^2 + \tfrac{1}{2}L_g\dot{\mathcal{Q}}^2 - \mathcal{Q}^2/2C_g + \tfrac{1}{2}L\dot{Q}^2 + (Q^2/2s)[1-(\mathcal{Q}/\mathcal{Q}_0)^2] - 2KxQ(\mathcal{Q}/\mathcal{Q}_0) + Kx(\dot{Q}/\omega)(\dot{\mathcal{Q}}/\omega\mathcal{Q}_0). \tag{B14}$$

A slightly prettier form can be obtained by the change of generator coordinate

$$\mathcal{Q}_{old} = \mathcal{Q}_{new} - (K/\mathcal{Q}_0 L_g\omega^2)xQ \tag{B15}$$

—a change which becomes $\mathcal{Q}_{old} = \mathcal{Q}_{new}$ in the limit $\varepsilon \to 0$ $(\mathcal{Q}_0/C_g \to \infty)$. By making this change of coordinate, and by discarding terms in $\mathcal{L}$ which vanish as $\varepsilon \to 0$, we bring our Lagrangian into the final form

$$\mathcal{L} = \tfrac{1}{2}m\dot{x}^2 - \tfrac{1}{2}m\omega^2 x^2 + \tfrac{1}{2}L_g\dot{\mathcal{Q}}^2 - \mathcal{Q}^2/2C_g + \tfrac{1}{2}L\dot{Q}^2 + (Q^2/2s)[1-(\mathcal{Q}/\mathcal{Q}_0)^2] - KQ[x(\mathcal{Q}/\mathcal{Q}_0) + (\dot{x}/\omega)(\dot{\mathcal{Q}}/\omega\mathcal{Q}_0)]. \tag{B16}$$

We next introduce the generalized momenta

$$p = \frac{\partial\mathcal{L}}{\partial\dot{x}} = m\dot{x} - \frac{KQ}{\omega}\left(\frac{\dot{\mathcal{Q}}}{\omega\mathcal{Q}_0}\right), \quad \Pi = \frac{\partial\mathcal{L}}{\partial\dot{Q}} = L\dot{Q}, \quad \mathcal{J} = \frac{\partial\mathcal{L}}{\partial\dot{\mathcal{Q}}} = L_g\dot{\mathcal{Q}} - \frac{KQ\dot{x}}{\omega^2\mathcal{Q}_0}, \tag{B17}$$

compute the Hamiltonian $H = p\dot{x} + \Pi\dot{Q} + \mathcal{J}\dot{\mathcal{Q}} - \mathcal{L}$, discard

terms that vanish as $\varepsilon \to 0$, and quantize. The result is

$$\hat{H} = \frac{\hat{p}^2}{2m} + \tfrac{1}{2}m\omega^2\hat{x}^2 + \frac{\hat{\Pi}^2}{2L} + \frac{\hat{Q}^2}{2s}\left[\left(\frac{\hat{\mathcal{J}}}{L_g\omega\mathcal{Q}_0}\right)^2 + \left(\frac{\hat{\mathcal{Q}}}{\mathcal{Q}_0}\right)^2 - 1\right] + \frac{\hat{\mathcal{J}}^2}{2L_g} + \frac{\hat{\mathcal{Q}}^2}{2C_g} + K\hat{Q}\left(\hat{x}\frac{\hat{\mathcal{Q}}}{\mathcal{Q}_0} + \frac{\hat{p}}{m\omega}\frac{\hat{\mathcal{J}}}{L_g\omega\mathcal{Q}_0}\right). \tag{B18}$$

c. Quantum generator compared with classical generator

Equation (B18) is the Hamiltonian of our oscillator plus measuring apparatus, with the generator treated quantum mechanically. Unruh (1979) has pointed out the importance of treating the generator quantum mechanically rather than classically in any fully rigorous analysis of measurements of X_1, and he was the first person to write Hamiltonian (B18) for such a fully rigorous analysis.

We now show that the quantum generator can be replaced, in principle, by a classical generator without loss of accuracy in our analysis—thereby justifying our use of classical generators throughout the text of this paper. Specifically, before any measurements begin the quantum generator is prepared in the coherent state of Eq. (B3), which has a mean number of quanta N_0, and has

$$\begin{aligned}
&\langle\hat{\mathcal{Q}}\rangle/\mathcal{Q}_0 = \cos\omega t, \quad \langle\hat{\mathcal{J}}\rangle/L_g\omega\mathcal{Q}_0 = -\sin\omega t,\\
&\Delta\mathcal{Q}/\mathcal{Q}_0 = \Delta\mathcal{J}/L_g\omega\mathcal{Q}_0 = \tfrac{1}{2}N_0^{-1/2},\\
&\langle\hat{\mathcal{J}}^2/2L_g + \hat{\mathcal{Q}}^2/2C_g\rangle = (N_0+\tfrac{1}{2})\hbar\omega,\\
&\langle(\hat{\mathcal{J}}/L_g\omega\mathcal{Q}_0)^2 + (\hat{\mathcal{Q}}/\mathcal{Q}_0)^2 - 1\rangle = \tfrac{1}{2}N_0^{-1},\\
&\Delta[(\mathcal{J}/L_g\omega\mathcal{Q}_0)^2 + (\mathcal{Q}/\mathcal{Q}_0)^2 - 1] = N_0^{-1/2}.
\end{aligned} \tag{B19}$$

Comparison of Eqs. (B19) and (B18) shows that, in the limit $N_0 \to \infty$ (i.e., $\varepsilon \to 0$), the generator behaves completely classically and is not loaded at all by the rest of the system—i.e., it is governed by the uncoupled Hamiltonian

$$\hat{H}_g = \hat{\mathcal{J}}^2/2L_g + \hat{\mathcal{Q}}^2/2C_g; \tag{B20}$$

and it always remains in the infinitely-sharply-peaked coherent state of Eq. (B19). The Hamiltonian for the rest of the system, when $N_0 \to \infty$, is obtained by removing the decoupled generator Hamiltonian (B20) from Eq. (B18), and by replacing $\hat{\mathcal{Q}}/\mathcal{Q}_0$ and $\hat{\mathcal{J}}/L_g\omega\mathcal{Q}_0$ by their sharp classical values, $\cos\omega t$ and $-\sin\omega t$. The result,

$$\hat{H} = \hat{p}^2/2m + \tfrac{1}{2}m\omega^2\hat{x}^2 + \hat{\Pi}^2/2L + K\hat{Q}[\hat{x}\cos\omega t - (\hat{p}/m\omega)\sin\omega t], \tag{B21}$$

is identical to Hamiltonian (3.16) with classical generator, which was analyzed in Sec. III. B of the text.

Suppose that the generator is not fully classical, i.e., that N_0 is finite. Then to what extent will a measurement of $\hat{X}_1$ be marred by quantum fluctuations in the generator and by loading of the generator by the experimental apparatus? The answer, when the exact Hamiltonian has the form of Eq. (B18), can be computed by a perturbation-theory analysis of the *Gedanken* experiment of Sec. III.B. Such a computation reveals the following, for the case where one wishes to measure $\hat{X}_1$ with accuracy better than the standard quantum limit $(\hbar/2m\omega)^{1/2}$ and with measurement time τ: Let μ be the fractional distance below the standard quantum limit which the experiment could achieve with a perfect, classical generator:

$$\mu \equiv \left(\frac{4\hbar L}{K^2\tau^3}\right)^{1/2}\left(\frac{1}{\hbar/2m\omega}\right)^{1/2} = \left(\frac{8mL\omega}{K^2\tau^3}\right)^{1/2} \tag{B22}$$

[cf. Eq. (3.21)]; if the probability distribution of $\hat{X}_1$ before the experiment begins is peaked about a value ξ_0 near zero, with variance $\Sigma = \alpha(\hbar/2m\omega)^{1/2}$ where $\alpha \lesssim \mu$, then the measurement (i) can determine ξ_0 with a probable error

$$\Delta\xi_m \simeq \mu\left(\frac{\hbar}{2m\omega}\right)^{1/2}\left[1 + \frac{\alpha^2}{\mu^2} + \text{terms of order}\left(\frac{1}{\mu^4(\omega\tau)^2N_0}, \frac{1}{\mu^4(\omega\tau)N_0}, \frac{1}{\mu^2\alpha^2N_0}\right)\right]^{1/2}, \tag{B23a}$$

and (ii) increases the variance of $\hat{X}_1$ to

$$\Delta X_1 \simeq \alpha\left(\frac{\hbar}{2m\omega}\right)^{1/2}\left[1 + \text{terms of order}\left(\frac{1}{\mu^4(\omega\tau)N_0}, \frac{1}{\mu^2\alpha^2N_0}\right)\right]^{1/2}. \tag{B23b}$$

Evidently the quantum properties of the generator cause negligible error in the experiment if the generator is excited in a coherent state with mean number of quanta

$$\langle\hat{N}\rangle \equiv N_0 \gg \max\left(\frac{1}{\mu^4(\omega\tau)^2}, \frac{1}{\mu^2\alpha^2}\right). \tag{B24}$$

Note that for measurements near the standard quantum limit ($\mu \sim \alpha \sim 1$) in times not much shorter than one cycle ($\omega\tau \gtrsim 1$), the generator does not need to be highly excited.

Unruh (1979) pointed out that one can design a quantum-mechanical generator which is protected entirely from loading (back action) by the experimental apparatus, even when the level of generator excitation is finite. To achieve such a "loading-free" generator one uses not a harmonic oscillator ($\hat{H}_g = \hat{\mathcal{J}}^2/2L_g + \hat{\mathcal{Q}}^2/2C_g$; $[\hat{\mathcal{Q}}, \hat{\mathcal{J}}] = i\hbar$), but rather the following system with two dynamical degrees of freedom:

$$\hat{H}_g = \frac{\hat{\mathcal{J}}^2}{2L_g} + \tfrac{1}{2}L_g\omega^2\hat{\mathcal{Q}}^2 - \frac{\hat{q}\hat{\mathcal{J}}}{L_g} + L_g\omega^2\hat{j}\hat{\mathcal{Q}}, \tag{B25a}$$

$$[\hat{q}, \hat{\mathcal{Q}}] = i\hbar, \quad [\hat{j}, \hat{\mathcal{J}}] = i\hbar, \tag{B25b}$$

all other commutators vanish. Equation (B25a) is the Hamiltonian of a charged particle in a suitable constant magnetic field with a quadrupole electric field to cancel the quadratic $\hat{j}^2$ and $\hat{q}^2$ terms in the magnetic Hamiltonian [cf. the example between Eqs. (4.7′) and (4.9), with the change of notation $m \to 1/(L_g\omega^2)$, $\frac{1}{2}eB_0 \to 1/(L_g\omega)$, $\hat{x} \to \hat{q}$, $\hat{p}_x \to \hat{\mathcal{Q}}$, $\hat{y} \to L_g\omega\hat{j}$, $\hat{p}_y \to \hat{\mathcal{J}}/L_g\omega$]. If such a generator is used in our *Gedanken* experiment [replace $\hat{\mathcal{J}}^2/2L_g + \hat{\mathcal{Q}}^2/2C_g$ in (B18) by (B25a)], the resulting Heisenberg

equations for $\hat{\mathcal{J}}$ and $\hat{\mathcal{Q}}$ will be precisely those of a free harmonic oscillator:

$$d\hat{\mathcal{J}}/dt = -L_g\omega^2\hat{\mathcal{Q}}, \quad d\hat{\mathcal{Q}}/dt = \hat{\mathcal{J}}/L_g. \tag{B26}$$

This shows the complete absence of loading of Unruh's generator by our experiment, independent of the state of the generator. However, quantum fluctuations (uncertainties) are still present in Unruh's generator and can affect the experiment—unless one puts Unruh's generator into a state with arbitrarily small variances of $\hat{\mathcal{Q}}$ and $\hat{\mathcal{J}}$ (possible because $[\hat{\mathcal{Q}}, \hat{\mathcal{J}}] = 0$). Such a special state

$$\langle\hat{\mathcal{Q}}\rangle = \mathcal{Q}_0\cos\omega t, \quad \langle\hat{\mathcal{J}}\rangle = -L_g\omega\mathcal{Q}_0\sin\omega t, \qquad \Delta\mathcal{J} = L_g\omega\Delta\mathcal{Q} = \text{const} \to 0, \tag{B27}$$

is the analog of the arbitrarily energetic coherent state ($\mathcal{Q}_0 \to \infty$) which our generator requires in order to avoid quantum fluctuations. Our generator's coherent state has an arbitrarily large expectation value and variance for its energy:

$$\langle\hat{H}_g\rangle = (N_0 + \tfrac{1}{2})\hbar\omega \to \infty,$$
$$\Delta H_g = N_0^{1/2}\hbar\omega = (\mathcal{Q}_0/2\Delta\mathcal{Q})\hbar\omega \to \infty. \tag{B28a}$$

Unruh's special state (B27), in principle, can have a finite mean energy $\langle\hat{H}_g\rangle$; but its energy variance is arbitrarily large and, in fact, for given $\mathcal{Q}_0$ and $\Delta\mathcal{Q}$ is of the same magnitude as our variance:

$$\Delta H_g\Delta\mathcal{Q} \geq \tfrac{1}{2}\left|\langle[\hat{\mathcal{Q}}, \hat{H}_g]\rangle\right| = (\hbar/2L_g)\left|\langle\hat{\mathcal{J}}\rangle\right| = \tfrac{1}{2}\hbar\omega\mathcal{Q}_0\left|\sin\omega t\right|,$$
$$\Delta H_g\Delta\mathcal{J} \geq \tfrac{1}{2}\left|\langle[\hat{\mathcal{J}}, \hat{H}_g]\rangle\right| = \tfrac{1}{2}\hbar L_g\omega^2\left|\langle\hat{\mathcal{Q}}\rangle\right| = \tfrac{1}{2}\hbar\omega(L_g\omega\mathcal{Q}_0)\left|\cos\omega t\right|,$$

whence

$$\Delta H_g \geq 2^{-1/2}(\mathcal{Q}_0/2\Delta\mathcal{Q})\hbar\omega \to \infty. \tag{B28b}$$

Unruh's generator [charged-particle system described by Eq. (B25)] was mentioned in Sec. IV.A as an illustration of the concept of a "generalized QND observable." The observables $\hat{\mathcal{Q}}$ and $\hat{\mathcal{J}}$ are a pair of such observables, and it is precisely this fact that allows Unruh's generator to avoid back action (loading) from the experiment.

Unruh's generator is important because it shows that in principle one can design a generator which is completely free of back action. However, it is not clear how one could realize physically the desired coupling of Unruh's generator to our experiment.

2. Electromagnetic oscillator

We now turn to a physical realization of Hamiltonian (3.16) for the case of an electromagnetic oscillator. Such a realization was given in *Physical Review Letters* (Thorne *et al.*, 1978) and is reproduced with minor changes in Fig. 10.

The oscillator whose X_1 is to be measured is an "LC circuit" consisting of the two coils (total self-inductance m) near the bottom of Fig. 10, and the four capacitor plates A, A', B, B' near the top. The oscillator is coupled, via coordinate (charge x) and momentum

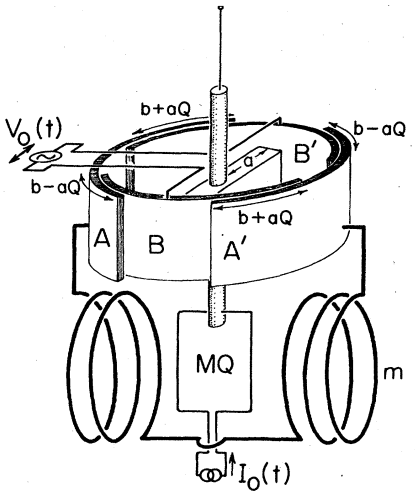

FIG. 10. Idealized physical realization of a system for measuring the $\hat{X}_1$ of an electromagnetic oscillator arbitrarily quickly and accurately. See text for discussion.

(magnetic flux p) transducers to a torsion pendulum (vertical central rod in Fig. 10, and paraphernalia attached to it). The coupling produces a torque $-KX_1$ on the torsion pendulum, causing it to swing through an angle Q. The coupling to $X_1 = x\cos\omega t - (p/m\omega)\sin\omega t$ requires a sinusoidal voltage $V_0(t) \propto \cos\omega t$ in the coordinate (charge x) transducer, and a sinusoidal current $I_0(t) \propto \sin\omega t$ in the momentum (magnetic flux p) transducer. The sinusoidal voltage and current are produced by an electromagnetic generator analogous to that in Sec. B.1, which is excited into an arbitrarily energetic coherent state. As sketched in Sec. B.1, this generator produces a perfect, classical output. For simplicity we here ignore its details and replace it by ideal classical voltage and current sources $V_0(t)$ and $I_0(t)$.

We now describe the *Gedanken* apparatus in greater detail. The *LC oscillator* (coils m and capacitor A-B-A'-B' in Fig. 10) is described mathematically by the charge x on plate A, the current $\dot{x}$ that flows through the coils, the total self-inductance m of the coils, the total capacitance C between plates A and A' (via B, B', and the zero-impedance voltage source connecting them), and the eigenfrequency $\omega = (1/mC)^{1/2}$ of the circuit's oscillations. The *coordinate (charge x) transducer* consists of plates B and B', to which are applied a sinusoidal voltage difference $V_0 \equiv -(b/a)K\cos\omega t$, and which are mechanically attached to the torsion pendulum. This voltage, together with the oscillator's charge x, produces a torque $\Gamma = -Kx\cos\omega t - (K^2/m\omega^2)Q\cos^2\omega t$ on the pendulum. *The velocity (current) transducer* consists of the thin wire loop at the bottom of Fig. 10, through which a sinusoidal current $I_0 \equiv (K/M\omega)\sin\omega t$ is driven. The loop is attached to the central rod so that its mutual inductance with the oscillator MQ is proportional to the angular displacement Q of the torsion pendulum. Current in the oscillator produces a torque $\Gamma = K(\dot{x}/\omega)\sin\omega t$ on the pendulum. The *torsion pendulum* (consisting of the central rod and paraphernalia attached to it and the torsion fiber that suspends

it) is characterized by its moment of inertia L, torsional spring constant $L\Omega^2$, natural frequency (in the absence of couplings) Ω, and generalized coordinate (equal to angular displacement) Q.

The complete apparatus—LC oscillator plus transducers plus torsion pendulum—is described by the classical Lagrangian

$$\mathcal{L} = \tfrac{1}{2}m(\dot{x}^2 - \omega^2 x^2) + \tfrac{1}{2}L(\dot{Q}^2 - \Omega^2 Q^2) - \tfrac{1}{2}(K^2/m\omega^2)(Q^2\cos^2\omega t) - K[x\cos\omega t - (\dot{x}/\omega)\sin\omega t]Q . \quad \text{(B29)}$$

The generalized momenta of the oscillator and pendulum are $p = \partial\mathcal{L}/\partial\dot{x} = m\dot{x} + (K/\omega)(\sin\omega t)Q$ and $\Pi = \partial L/\partial\dot{Q} = L\dot{Q}$; and the Hamiltonian, after quantization, is

$$\hat{H} = \hat{p}^2/2m + \tfrac{1}{2}m\omega^2\hat{x}^2 + K\hat{X}_1\hat{Q} + \hat{\Pi}^2/2L + \tfrac{1}{2}L\bar{\Omega}^2\hat{Q}^2 . \quad \text{(B30)}$$

Here the eigenfrequency $\bar{\Omega}$ of the pendulum is shifted from its natural value Ω by coupling to the coordinate and momentum sensors:

$$\bar{\Omega}^2 = \Omega^2 + K^2/m\omega^2 L . \quad \text{(B31)}$$

The frequency renormalization (B31) comes from two sources: *First*, the velocity (current) transducer used in the apparatus is equivalent to a momentum $[p = m\dot{x} + (KQ/\omega)\sin\omega t]$ transducer plus a positive spring on the torsion pendulum with spring constant $(K^2/m\omega^2)\sin^2\omega t$. (This is the analog, for measurements of electromagnetic oscillators, of our "velocity sensor equals momentum sensor plus positive capacitance" in Secs. III.A.1 and III.B.) *Second*, the "concentric-tin-can" shape of our capacitor-plus-coordinate-transducer (Fig. 10) is carefully designed to produce on the torsion pendulum a restoring torque with spring constant $(K^2/m\omega^2)\cos^2\omega t$. This was done so that the net renormalization of the pendulum's eigenfrequency would be time independent.

Hamiltonian (B30) will have the desired form [Eq. (3.16)] for a quick and accurate measurement of $\hat{X}_1$, if we set $\bar{\Omega}^2 = 0$. This requires that the natural eigenfrequency Ω of the torsion pendulum be imaginary

$$\Omega^2 = -K^2/m\omega^2 L , \quad \text{(B32)}$$

i.e., that the pendulum possess a noiseless spring with negative spring constant $-K^2/m\omega^2$. This negative spring is the analog of the negative capacitors needed in Secs. III.A.2 and B.1; cf. also footnote 11 in Appendix A. Figure 11 shows an idealized example of such a noiseless negative spring.

APPENDIX C: ARBITRARILY QUICK AND ACCURATE BACK-ACTION-EVADING MEASUREMENTS OF X_1: A DETAILED QUANTUM-MECHANICAL ANALYSIS[12]

1. Overview

This Appendix builds upon and expands the discussion given in Sec. III.B; the objective is to give a detailed quantum-mechanical analysis of a sequence of measurements of the X_1 of a harmonic oscillator. The analysis is exact quantum mechanically, and it should satisfy a theorist's desire for rigor. However, this rigor is purchased at the price of a highly idealized description of the measurement process, and this idealization may make an experimenter uneasy. He may prefer the more realistic, but semiclassical, measurement analyses given in Paper II.

Presenting two different analyses to appeal to two different constituencies may seem more like politics than physics, but we plead principle as well as pragmatism for the practice. We give an exact quantum-mechanical analysis of a simple, idealized version of a real measurement. We then ask whether a semiclassical treatment of a similar, simple system gives the same results. If it does, we gain the confidence to apply semiclassical techniques to complicated, realistic measuring systems—systems so complex that an exact quantum-mechanical treatment would be exceedingly difficult.

The key word in this Appendix is *sequence*. Section III.B of the body of this paper described apparatus for measuring X_1 and analyzed a single measurement of X_1 using this apparatus. The analysis proceeded by calculating the free evolution of the coupled oscillator-meter system [Eqs. (3.17)–(3.19)], and it demonstrated that X_1 can be measured arbitrarily quickly and arbitrarily accurately [Eq. (3.21)]. In this Appendix we string together a sequence of measurements of the type considered in Sec. III.B. To analyze the sequence, we must

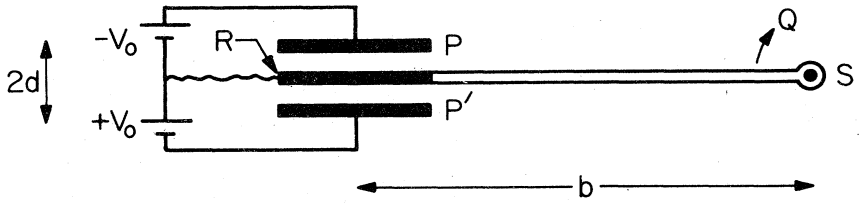

FIG. 11. Idealized example of a negative spring attached to a torsion pendulum. A dc bias voltage $-V_0$ is applied to the upper plate P, and a voltage $+V_0$ to the lower plate P', of a parallel-plate capacitor. The middle plate R is held at ground potential by the wavy wire, and is physically attached by a lever arm of length b to the central shaft S of the torsion pendulum. When the shaft rotates through a small angle Q from equilibrium, plate R moves upward by a distance bQ; and the batteries V_0 drive a charge $q = (bQ/d)CV_0$ onto R. (Here $C = 2\mathcal{A}/4\pi d$ is the capacitance of plates $P - P'$ relative to plate R, and we assume $bQ \ll d$.) The charge q couples to the electric field V_0/d in the capacitor, producing an antirestoring torque $\Gamma = bqV_0/d = (b/d)^2 CV_0^2 Q$.

[12]The ideas and prose of this Appendix are due entirely to Carlton M. Caves, and constitute a portion of the material submitted by him to the California Institute of Technology in partial fulfillment of the requirements for the Ph.D. degree.

do more than just calculate the free quantum-mechanical evolution of the system; we must also have a rule for carrying the quantum state from one measurement to the next. That rule is the "reduction of the wave function" at the end of each measurement [see Eq. (C28)]. Free evolution and reduction of the wave function—together, these two allow us to follow changes in the state of the oscillator from measurement to measurement, and with this knowledge we can investigate the behavior of the oscillator during a sequence of measurements.

One important issue is the question of how X_1 changes during a sequence. Two results of our analysis bear on this issue. The first is that between measurements the expectation value of $\hat{X}_1$ can change, with the expected change always less than or equal to the variance of $\hat{X}_1$. The second is that the variance of $\hat{X}_1$ always decreases from one measurement to the next, for the type of measurement we analyze. Putting these two results together, we show that the expected change in the expectation value of $\hat{X}_1$ during a sequence of measurements is approximately the variance of $\hat{X}_1$ before the initial measurement.

Another, perhaps more important, issue is the question of how X_2 changes. In each measurement of the type in Sec. III.B, the expectation value of X_2 receives a large "kick" because the meter coordinate Q gets displaced a large distance from zero [cf. Eqs. (3.17)–(3.19)]. These kicks accumulate from one measurement to the next, and the expectation value of $\hat{X}_2$ runs away. However, these "expectation-value kicks" are essentially classical and predictable, so one might think that the resulting "classical runaway of X_2" could be avoided by applying a "feedback force" to the meter—a force whose purpose is to keep the meter coordinate close to zero. We investigate this issue using a model for the feedback, and we show that feedback can indeed keep the expectation value of $\hat{X}_2$ from running away. However, only part of each kick is classical. The feedback, no matter how good it may be, cannot eliminate the huge unpredictable quantum-mechanical kick given X_2 by each precise measurement of X_1, a kick whose size is determined by the uncertainty principle, Eq. (2.9a). One might expect these "uncertainty principle kicks" to add randomly, thereby causing X_2 to random walk. We verify the existence of this "random walk of X_2" by showing that, during a sequence of measurements, the variance of $\hat{X}_2$ increases as the square root of the number of measurements.

We choose to ignore the classical driving force $F(t)$ in this Appendix. Its effect on the oscillator could be included in the analysis. However, Sec. III.B has already shown that the classical force can be measured arbitrarily quickly and arbitrarily accurately. In addition, the classical force is irrelevant to the issues addressed in this Appendix. Its inclusion would only complicate the analysis without adding any new insights.

2. Description of the measuring apparatus

We now turn our attention to a detailed description of the measurement process. We begin by describing the physical system, which is nearly the same as that in Sec. III.B. The oscillator to be measured is characterized by the variables of Eqs. (2.1)–(2.5), including coordinate $\hat{x}$, momentum $\hat{p}$, complex amplitude $\hat{X}_1 + i\hat{X}_2$, mass m, and frequency ω. The oscillator is coupled to a measuring apparatus which consists of three parts: a generator, a meter, and a readout system. The generator provides the sinusoidal coupling in the interaction Hamiltonian. The meter is a one-dimensional quantum-mechanical "free mass" with generalized coordinate $\hat{Q}$, generalized momentum $\hat{\Pi}$, and generalized mass L; the meter is coupled by the generator to $\hat{X}_1$ of the oscillator. The readout system is coupled to the meter in such a way that at designated moments of time it "reads out" a value for the meter's coordinate Q and that at all times during each measurement it applies a constant "feedback force" to the meter. The feedback force is included to prevent the classical runaway of X_2.

Of the three parts of the measuring apparatus, only the meter will be treated quantum mechanically. As is discussed in Appendix B.1, the generator can be treated classically if, before the initial measurement in the sequence, it is prepared in a coherent state of arbitrarily large amplitude. Then the generator is completely unloaded by its coupling to the rest of the system, and it produces perfect "$\cos\omega t$" and "$\sin\omega t$" terms in the Hamiltonian.

The readout system will also be treated classically—i.e., we place the "quantum-classical cut" of our analysis between the meter and the readout system. This choice is legitimate if inclusion of all or part of the readout system in the quantum-mechanical analysis would not substantially degrade the calculated accuracy of the measurement. For example, the readout system can in principle be a device which is so strongly coupled to the meter that it makes arbitrarily precise, essentially instantaneous measurements of the meter coordinate (see discussion in footnote 6). This is the model we shall adopt. Then a "readout" of Q by the readout system is described as follows: The readout system determines a value for the meter coordinate Q at a particular instant, thereby localizing the meter precisely at the measured value; formally this means that the quantum state of the oscillator-meter system is "reduced" to an eigenstate of $\hat{Q}$ whose eigenvalue is the measured value [see Eq. (C28)].

In practice the readout system will not make infinitely precise measurements of the meter coordinate. We shall consider the case of a finite-precision readout system in Sec. C.7, where we sketch a density-matrix analysis of a sequence of measurements of X_1.

Finally, we also treat the feedback force classically.

3. Foundations for the analysis

The Hamiltonian for the coupled oscillator-meter system has the form of Eq. (3.16) with the addition of a term describing the feedback force:

$$\hat{H} = \hat{H}_0 + \hat{H}_M + \hat{H}_I , \tag{C1a}$$

$$\hat{H}_0 = \hat{p}^2/2m + \tfrac{1}{2}m\omega^2\hat{x}^2 , \tag{C1b}$$

$$\hat{H}_M = \hat{\Pi}^2/2L , \tag{C1c}$$

$$\hat{H}_I = K(\hat{X}_1 - \alpha)\hat{Q} . \tag{C1d}$$

Here $\hat{H}_0$ is the Hamiltonian of the free oscillator, $\hat{H}_M$ is the Hamiltonian of the meter, and K is a coupling constant. The interaction Hamiltonian $\hat{H}_I$ consists of two terms: a term $K\hat{X}_1\hat{Q}$ which describes the coupling of the oscillator to the meter via the classical generator, and a term $-K\alpha\hat{Q}$ which describes the classical feedback force on the meter. The size of the feedback force is determined by the parameter α ("force"$=K\alpha$). The feedback is under the control of the experimenter; in general, α will change from measurement to measurement in the sequence. Designs for physical systems which are governed in principle by the Hamiltonian (C1) are considered in Appendix B. Here we do not concern ourselves with any specific physical system. The Hamiltonian (C1) is the starting point of the analysis, which applies to any system governed by that Hamiltonian.

The analysis in Sec. III.B uses the Heisenberg picture. It is the most convenient picture for calculating the evolution of the expectation value and variance of $\hat{Q}$ [Eqs. (3.19)], and those are the only results really necessary for that analysis. In this Appendix we work exclusively in the Schrödinger picture. This is not because the Heisenberg picture could not be used; rather, it is because the Schrödinger picture is more convenient and more natural for analyzing a sequence of measurements. In particular, the reduction of the wave function can be handled more easily in the Schrödinger picture.

In the Heisenberg picture the complex amplitude of a free harmonic oscillator is conserved [Eq. (2.7)]. In the Schrödinger picture the operators $\hat{X}_1$ and $\hat{X}_2$ are time dependent, and whenever it is necessary, we shall indicate explicitly the time at which they are evaluated:

$$\hat{X}_1(t)=\hat{x}\cos\omega t-(\hat{p}/m\omega)\sin\omega t\,, \tag{C2a}$$

$$\hat{X}_2(t)=\hat{x}\sin\omega t+(\hat{p}/m\omega)\cos\omega t\,. \tag{C2b}$$

The corresponding Heisenberg operators for a free harmonic oscillator are given by

$$(\hat{X}_j)_H(t)\equiv\hat{U}_0^\dagger(t,t_0)\hat{X}_j(t)\hat{U}_0(t,t_0)\,,$$

where $\hat{U}_0$ is the time-development operator for the free oscillator:

$$\hat{U}_0(t,t_0)=e^{-i(t-t_0)\hat{H}_0/\hbar}\,. \tag{C3}$$

Hence, conservation of the complex amplitude of a free oscillator translates into the following identity in the Schrödinger picture:

$$\hat{X}_j(t)\equiv\hat{U}_0(t,t_0)\hat{X}_j(t_0)\hat{U}_0^\dagger(t,t_0)\,. \tag{C4}$$

Equation (C4) holds for arbitrary times t and t_0.

In the Schrödinger picture the information about the state vector $|\psi(t)\rangle$ of the coupled oscillator-meter system is conveniently expressed in terms of an evolving "wave function," which is defined by projecting $|\psi(t)\rangle$ onto appropriate basis states. For the meter the choice of basis states is obvious. Since we are interested in the behavior of the meter coordinate, we choose the eigenstates $|Q\rangle$ of $\hat{Q}$ with delta-function normalization:

$$\hat{Q}|Q\rangle=Q|Q\rangle\,,\quad \langle Q|Q'\rangle=\delta(Q-Q')\,. \tag{C5}$$

For the oscillator the most convenient basis states are eigenstates of $\hat{X}_1(t)$. To define such states we begin with the delta-function normalized eigenstates $|\xi,0\rangle$ of $\hat{X}_1(0)=\hat{x}$:

$$\hat{X}_1(0)|\xi,0\rangle=\xi|\xi,0\rangle\,,\quad \langle\xi,0|\xi',0\rangle=\delta(\xi-\xi')\,. \tag{C6}$$

We then define new states

$$|\xi,t\rangle\equiv\hat{U}_0(t,0)|\xi,0\rangle\,. \tag{C7}$$

These new states have delta-function normalization, and as one shows using Eq. (C4), they are also the desired eigenstates of $\hat{X}_1(t)$:

$$\hat{X}_1(t)|\xi,t\rangle=\xi|\xi,t\rangle\,,\quad \langle\xi,t|\xi',t\rangle=\delta(\xi-\xi')\,. \tag{C8}$$

An important property of these states is that

$$|\xi,t\rangle=\hat{U}_0(t,t_0)|\xi,t_0\rangle\,. \tag{C9}$$

A complete set of states for the oscillator-meter system can be obtained by taking the tensor product of the states $|\xi,t\rangle$ and $|Q\rangle$:

$$|\xi,Q;t\rangle\equiv|\xi,t\rangle\otimes|Q\rangle\,. \tag{C10}$$

Given this complete set, we can define a wave function corresponding to the state vector $|\psi(t)\rangle$ by

$$\psi(\xi,Q;t)\equiv\langle\xi,Q;t|\psi(t)\rangle\,. \tag{C11}$$

The wave function has the usual interpretation: $|\psi(\xi,Q;t)|^2\,d\xi\,dQ$ is the probability at time t of simultaneously finding the meter coordinate between Q and $Q+dQ$ and the oscillator with X_1 between ξ and $\xi+d\xi$.

In the Schrödinger picture the evolution of the state vector $|\psi(t)\rangle$ is determined by the unitary time-development operator $\hat{U}(t,t_0)$ [not to be confused with $\hat{U}_0(t,t_0)$]—i.e.,

$$|\psi(t)\rangle=\hat{U}(t,t_0)|\psi(t_0)\rangle\,, \tag{C12a}$$

where $\hat{U}(t,t_0)$ satisfies the Schrödinger equation

$$i\hbar\,d\hat{U}(t,t_0)/dt=\hat{H}(t)\hat{U}(t,t_0)\,,\quad \hat{U}(t_0,t_0)=1\,. \tag{C12b}$$

For Hamiltonian (C1), the solution for $\hat{U}(t,t_0)$ can be obtained using the techniques employed to solve for the time-development operator of a forced harmonic oscillator [Eqs. (2.20)]. We omit the details and simply give the solution:

$$\begin{aligned}\hat{U}(t,t_0)=\hat{U}_0(t,t_0)&\exp\left(-\frac{i}{\hbar}(t-t_0)^3\frac{K^2[\hat{X}_1(t_0)-\alpha]^2}{6L}\right)\\&\times\exp\left(-\frac{i}{\hbar}(t-t_0)K[\hat{X}_1(t_0)-\alpha]\hat{Q}\right)\\&\times\exp\left(\frac{i}{\hbar}(t-t_0)^2\frac{K[\hat{X}_1(t_0)-\alpha]}{2L}\hat{\Pi}\right)\hat{U}_M(t,t_0)\,,\end{aligned} \tag{C13}$$

where

$$\hat{U}_M(t,t_0)=e^{-i(t-t_0)\hat{H}_M/\hbar} \tag{C14}$$

is the time-development operator for the meter Hamiltonian.

The abstract operator equations (C12) and (C13) governing the evolution of the state vector can be translated into an equivalent equation for the evolution of the wave function. In the case of interest to us the oscillator and meter are in states $|\chi\rangle$ and $|\Phi\rangle$, respectively,

at time t_0; the corresponding wave functions are $\chi(\xi) \equiv \langle \xi, t_0 | \chi \rangle$ and $\Phi(Q) \equiv \langle Q | \Phi \rangle$. The initial state vector is $|\psi(t_0)\rangle = |\chi\rangle \otimes |\Phi\rangle$, with associated wave function $\psi(\xi, Q; t_0) = \chi(\xi)\Phi(Q)$. Since $e^{-i\alpha\hat{\Pi}/\hbar}$ is the displacement operator for the meter (see, e.g., Sec. 14.7 of Merzbacher, 1970), it is easy to show that

$$e^{-i\nu[\hat{X}_1(t_0)-\alpha]\hat{\Pi}/\hbar} |\xi, Q; t_0\rangle = |\xi, Q + \nu(\xi - \alpha); t_0\rangle , \qquad \text{(C15)}$$

where ν is any real number. Using Eqs. (C5), (C8)–(C13), and (C15), one can derive the following equation for the evolution of the wave function:

$$\psi(\xi, Q; t) = \exp\left(-\frac{i}{\hbar}(t-t_0)^3 \frac{K^2(\xi-\alpha)^2}{6L}\right) \times \exp\left(-\frac{i}{\hbar}(t-t_0)K(\xi-\alpha)Q\right) \times \chi(\xi)\Phi_{\text{free}}\left(Q + (t-t_0)^2\frac{K(\xi-\alpha)}{2L}, t\right). \qquad \text{(C16a)}$$

Here $\Phi_{\text{free}}(Q, t)$ is the wave function which gives the evolution of a "free mass" whose initial state is $|\Phi\rangle$, i.e.,

$$\Phi_{\text{free}}(Q, t) \equiv \int dQ' \mathcal{K}_{\text{free}}(Q, Q'; t, t_0)\Phi(Q') , \qquad \text{(C16b)}$$

where

$$\mathcal{K}_{\text{free}}(Q, Q'; t, t_0) \equiv \langle Q | \hat{U}_M(t, t_0) | Q' \rangle \qquad \text{(C16c)}$$

is the kernel of the free-mass Schrödinger equation. The explicit form of $\mathcal{K}_{\text{free}}$ is given in many standard quantum mechanics text books [see, e.g., Eq. (8.91) of Merzbacher (1970)].

The wave function (C16) shows particularly clearly the effect of the interaction on the meter. The probability distribution of the meter coordinate at time t is

$$P(Q) = \int |\psi(\xi, Q; t)|^2 d\xi = \int P(Q|\xi)|\chi(\xi)|^2 d\xi , \qquad \text{(C17a)}$$

$$P(Q|\xi) \equiv \left|\Phi_{\text{free}}\left(Q + (t-t_0)^2\frac{K(\xi-\alpha)}{2L}, t\right)\right|^2 . \qquad \text{(C17b)}$$

Here $P(Q|\xi)$ can be regarded as a conditional probability distribution—i.e., the probability distribution of Q given that $X_1 = \xi$. The important feature of $P(Q|\xi)$ is this: It has the same shape as the probability distribution for a "free mass," but it is displaced a distance $-K(\xi-\alpha)(t-t_0)^2/2L$—precisely the displacement produced by a classical force $-K(\xi-\alpha)$.

4. Analysis of a single measurement

We are now ready to analyze a single measurement in detail—the first task in constructing a sequence of measurements. For generality we let the particular measurement under consideration be the nth in the sequence. The measurement process can be described in general terms as follows. Before the nth measurement the meter is prepared in an appropriate initial state, and the oscillator is in some state left over from the preceding measurement. At time t_{n-1} the interaction is turned on, and the oscillator and meter are allowed to interact freely for a time τ. At time $t_n = t_{n-1} + \tau$ the interaction is turned off, the readout system makes an infinitely precise "measurement" of the meter coordinate, and the wave function is reduced. (We shall call this precise "measurement" of Q a "readout" to avoid confusion with the "measurement of X_1" which lasts from t_{n-1} to t_n.) The reduction of the wave function is the link that connects this measurement to the next one. It allows us to identify the state of the oscillator *after* the measurement—a state which becomes the initial oscillator state for the next measurement.

Two aspects of this process deserve special attention. The first is that the oscillator-meter coupling is on only during the interval from t_{n-1} to t_n: The interaction is turned on abruptly at time t_{n-1} and turned off abruptly at time t_n (functional form of K for nth measurement: $K = 0$ for $t < t_{n-1}$ and $t > t_n$, $K = \text{const} \neq 0$ for $t_{n-1} < t < t_n$). The step-function form of K is not the important issue; less abrupt forms for K could be used without changing the results significantly. The important point is that preparation of the meter is done with the interaction turned off. In a real experiment one would probably leave the interaction on while the meter is prepared. One could do so without affecting X_1, because X_1 is completely isolated from the meter; however, X_2 would be affected [cf. Eqs. (3.17)]. Since one of our objectives is to investigate the behavior of X_2, we choose to prepare the meter with the interaction turned off. Then X_2 is unaffected by meter preparation. Indeed, while the interaction is turned off, the oscillator's X_1 wave function is constant.

The second important aspect is that we regard each measurement in the sequence as beginning at the instant when a readout terminates the preceding measurement. This is purely a matter of convenience. If the reader wishes to insert a time interval between measurements to allow for meter preparation or any other activity, she can do so. Our results will not be affected, because the oscillator's X_1 wave function is constant while the interaction is turned off.

All quantities characteristic of the time interval $t_{n-1} \leq t < t_n$ will be denoted by a subscript $n-1$—e.g., state vector $|\psi_{n-1}(t)\rangle$, wave function $\psi_{n-1}(\xi, Q; t)$, feedback parameter α_{n-1}. The values measured at time t_n will be denoted by a subscript n.

We now consider in turn each of the four components of the nth measurement: specification of the initial state, free evolution of the coupled oscillator-meter, readout of the meter coordinate, and reduction of the wave function.

At time t_{n-1} the oscillator is in some state $|\chi_{n-1}\rangle$ with wave function $\chi_{n-1}(\xi) = \langle \xi, t_{n-1} | \chi \rangle$; except for the first measurement, this state is left over from the previous measurement. The associated expectation value and variance of $\hat{X}_1(t_{n-1})$ we denote $\langle X_1 \rangle_{n-1}$ and $(\Delta X_1)_{n-1}$; similarly, for $\hat{X}_2(t_{n-1})$, $\langle X_2 \rangle_{n-1}$ and $(\Delta X_2)_{n-1}$. The meter is prepared in a Gaussian (minimum-uncertainty) wave-packet state $|\Phi\rangle$ with wave function

$$\Phi(Q) = [2\pi(\Delta Q)_0^2]^{-1/4} \exp[-Q^2/4(\Delta Q)_0^2] . \qquad \text{(C18)}$$

This state has $\langle \hat{Q} \rangle = \langle \hat{\Pi} \rangle = 0$. We choose the variances $(\Delta Q)_0 = (\hbar\tau/2L)^{1/2}$, $(\Delta\Pi)_0 = (\hbar L/2\tau)^{1/2}$—a choice which minimizes the variance of $\hat{Q}$ at time t_n. The initial state vector is $|\psi_{n-1}(t_{n-1})\rangle = |\chi_{n-1}\rangle \otimes |\Phi\rangle$, with wave function $\psi_{n-1}(\xi, Q; t_{n-1}) = \chi_{n-1}(\xi)\Phi(Q)$. Finally, the experimenter must also choose a value α_{n-1} for the feed-

back parameter.

The interaction is turned on at time t_{n-1}, and the coupled oscillator-meter evolves freely for a time τ. The evolution of the wave function during this interval can be obtained by specializing Eqs. (C16) to quantities characteristic of the nth measurement. Integration of Eq. (C16b) using the particular form of Eq. (C18) for $\Phi(Q)$ yields

$$\Phi_{\rm free}(Q,t)=[2\pi(\Delta Q)_0^2]^{-1/4}\left(1+i\frac{u}{\tau}\right)^{-1/2} \times\exp\left[-\frac{Q^2}{4(\Delta Q)_0^2}\left(1+i\frac{u}{\tau}\right)^{-1}\right], \tag{C19}$$

where $u\equiv t-t_{n-1}$. The effect of the interaction is to produce a strong correlation between the states of the meter and oscillator: At time $t_n=t_{n-1}+\tau$ the expectation value of the meter coordinate gets displaced to

$$\langle\hat{Q}\rangle(t_n)=-(K\tau^2/2L)(\langle X_1\rangle_{n-1}-\alpha_{n-1}), \tag{C20a}$$

and its variance becomes

$$\Delta Q(t_n)=[(\hbar\tau/L)+(K\tau^2/2L)^2(\Delta X_1)^2_{n-1}]^{1/2}. \tag{C20b}$$

Equations (C20) can be calculated directly from the probability distribution of the meter coordinate [Eqs. (C17) and (C19)] or, perhaps more easily, from a Heisenberg-picture analysis of the free evolution of the oscillator-meter system [cf. Eqs. (3.17)–(3.19)].

At time t_n the readout system reads out a value Q_n for the meter coordinate, and using Eq. (C20a) the experimenter infers a value

$$\xi_n=\alpha_{n-1}-(2L/K\tau^2)Q_n \tag{C21}$$

for X_1. The probability distribution of ξ_n, obtained directly from the probability distribution of Q [Eqs. (C17)], is given by

$$P(\xi_n)=(2\pi\sigma^2)^{-1/2}\int d\xi\,|\chi_{n-1}(\xi)|^2\exp\left(-\frac{(\xi-\xi_n)^2}{2\sigma^2}\right), \tag{C22}$$

where

$$\sigma\equiv(4\hbar L/K^2\tau^3)^{1/2}. \tag{C23}$$

This probability distribution refers to an ensemble of identical systems which begin the nth measurement in the same state. The mean and variance of ξ_n (averages over this ensemble) are

$$\bar{\xi}_n=\alpha_{n-1}-(2L/K\tau^2)\langle\hat{Q}\rangle(t_n)=\langle X_1\rangle_{n-1}, \tag{C24a}$$

$$\Delta\xi_n=[\sigma^2+(\Delta X_1)^2_{n-1}]^{1/2}. \tag{C24b}$$

Note that if $|\chi_{n-1}(\xi)|^2$ is a Gaussian, then $P(\xi_n)$ is also a Gaussian.

Equations (C24) tell us that the nth measurement can determine the expectation value $\langle X_1\rangle_{n-1}$ with probable error $\Delta\xi_n$. The error is minimized when $|\chi_{n-1}(\xi)|^2$ is highly peaked about its mean value [$(\Delta X_1)_{n-1}\ll\sigma$]; in this situation it makes sense to talk about X_1 having a particular value—a value which can be determined with error

$$\Delta\xi_n\simeq\sigma=(4\hbar L/K^2\tau^3)^{1/2}. \tag{C25}$$

Since $(\Delta X_1)_{n-1}$ can be arbitrarily small, *σ is the fundamental measure of the accuracy of X_1 measurements of duration τ, made with a meter of "mass" L which is perfectly coupled to X_1 with coupling constant K.* No matter how small τ may be, σ can be made as small as one wishes (in principle) by choosing K^2/L large enough; the measurements of X_1 can be arbitrarily quick and arbitrarily accurate.

This situation is to be contrasted with, for example, a measurement of the position of a free mass. There the feedback of momentum uncertainty onto position prevents a measurement of duration τ from having an accuracy better than $(\hbar\tau/m)^{1/2}$ [standard quantum limit for free-mass position; Eq. (3.2)].

A useful dimensionless characterization of the accuracy of X_1 measurement is provided by the ratio η of the standard quantum limit for amplitude-and-phase measurement [Eq. (2.16)] to σ:

$$\eta\equiv(1/\sigma)(\hbar/2m\omega)^{1/2},\quad \eta^2=K^2\tau^3/8mL\omega; \tag{C26}$$

η^{-1} is the factor by which measurements with given K, L, and τ beat this standard quantum limit.

When the readout determines a value Q_n for the meter coordinate at time t_n, it localizes the meter at $Q=Q_n$. This localization is described formally by projecting the state vector $|\psi_{n-1}(t_n)\rangle$ onto the eigenstate corresponding to the measured value (reduction of the wave function). We define a projection operator

$$\hat{\mathcal{P}}(Q)\equiv|Q\rangle\langle Q|=\int d\xi\,|\xi,Q;t_n\rangle\langle\xi,Q;t_n|, \tag{C27}$$

which projects the meter onto the eigenstate $|Q\rangle$. The state vector of the oscillator-meter system immediately after the readout is

$$|\psi_a\rangle=\mathfrak{A}\hat{\mathcal{P}}(Q_n)\,|\psi_{n-1}(t_n)\rangle =\left(\mathfrak{A}\int d\xi\,|\xi,t_n\rangle\psi_{n-1}(\xi,Q_n;t_n)\right)\otimes|Q_n\rangle \tag{C28}$$

[wave function $\psi_a(\xi,Q)=\mathfrak{A}\psi_{n-1}(\xi,Q_n;t_n)\delta(Q-Q_n)$], where $\mathfrak{A}$ is a normalization constant. ($\mathfrak{A}$ also contains an unknown, but irrelevant, phase factor which we shall ignore.) The state vector (C28) splits cleanly into oscillator and meter states. The oscillator state after the measurement becomes the initial state $|\chi_n\rangle$ for the $(n+1)$th measurement; its wave function is

$$\chi_n(\xi)=\mathfrak{A}\psi_{n-1}(\xi,Q_n;t_n). \tag{C29}$$

This wave function can be put in the form

$$\chi_n(\xi)=\mathfrak{B}\chi_{n-1}(\xi)\exp\{-(\xi^2-2\xi_n\xi)/4\sigma^2 +(i/\hbar)m\omega\eta^2[(3\xi_n-\tfrac{4}{3}\alpha_{n-1})\xi-\tfrac{5}{6}\xi^2]\} \tag{C30}$$

[Eqs. (C16), (C19), (C21), (C23), and (C26)], where $\mathfrak{B}$ is another normalization constant.

Equation (C30) is the fundamental equation of our analysis. It tells us how the oscillator wave function changes from one measurement to the next, and from it all our results will flow. One immediate consequence of Eq. (C30) is the following: If the oscillator begins the nth measurement in an eigenstate of $\hat{X}_1$ [$|\chi_{n-1}\rangle=|\xi',t\rangle$, $\chi_{n-1}(\xi)=\delta(\xi-\xi')$], then it remains in an eigenstate of $\hat{X}_1$ with the same eigenvalue after the measurement. As is discussed in Sec. IV, this is the essential feature of quantum nondemolition measurement.

5. Analysis of a sequence of measurements

Having completed our analysis of a single measurement, we turn next to analyzing a sequence of measurements and, in particular, to investigating the behavior of X_1 and X_2 during a sequence of measurements. To do so requires specifying a particular form for $\chi_{n-1}(\xi)$. The form we choose is

$$\chi_{n-1}(\xi)=[2\pi(\Delta X_1)^2_{n-1}]^{-1/4}\exp\left(-\frac{(\xi-\langle X_1\rangle_{n-1})^2}{4(\Delta X_1)^2_{n-1}}+\frac{i}{\hbar}m\omega(a_{n-1}\xi-\mu_{n-1}\xi^2)\right), \quad \text{(C31)}$$

where a_{n-1} and μ_{n-1} are real constants. If $\mu_{n-1}=0$, $\chi_{n-1}(\xi)$ is a minimum-uncertainty wave packet. Using the fact that in the ξ representation $\hat{X}_2$ is equivalent to $(\hbar/im\omega)(\partial/\partial\xi)$, one can readily evaluate the expectation value and variance of $\hat{X}_2(t_{n-1})$ associated with the wave function (C31):

$$\langle X_2\rangle_{n-1}=a_{n-1}-2\mu_{n-1}\langle X_1\rangle_{n-1}, \quad \text{(C32a)}$$

$$(\Delta X_2)^2_{n-1}=(\hbar/2m\omega)^2(\Delta X_1)^{-2}_{n-1}+4\mu^2_{n-1}(\Delta X_1)^2_{n-1}. \quad \text{(C32b)}$$

The reason for the choice (C31) should be clear. As a glance at Eq. (C30) shows, the form (C31) for the initial oscillator wave function is preserved from one measurement to the next in a sequence; the only things that change are the constants characterizing the wave function:

$$(\Delta X_1)^{-2}_n=(\Delta X_1)^{-2}_{n-1}+\sigma^{-2}, \quad \text{(C33a)}$$

$$\frac{\langle X_1\rangle_n}{(\Delta X_1)^2_n}=\frac{\langle X_1\rangle_{n-1}}{(\Delta X_1)^2_{n-1}}+\frac{\xi_n}{\sigma^2}, \quad \text{(C33b)}$$

$$a_n=a_{n-1}+\eta^2(3\xi_n-\tfrac{4}{3}\alpha_{n-1}), \quad \text{(C33c)}$$

$$\mu_n=\mu_{n-1}+\tfrac{5}{6}\eta^2. \quad \text{(C33d)}$$

The first of these equations has a couple of immediate consequences. The first is that $(\Delta X_1)_n\leq(\Delta X_1)_{n-1}$; hence, as a sequence of measurements proceeds, $|\chi_n(\xi)|^2$ becomes more and more highly peaked. The second is that if the oscillator is in a state with $\Delta X_1\gg\sigma$, one measurement is sufficient to prepare it in a state with $\Delta X_1\simeq\sigma$.

By manipulating Eqs. (C33) with the help of Eqs. (C24), one can show that the change in the expectation value of $\hat{X}_1$ in the nth measurement is

$$\langle X_1\rangle_n-\langle X_1\rangle_{n-1}=[(\Delta X_1)_{n-1}/\Delta\xi_n]^2(\xi_n-\bar{\xi}_n). \quad \text{(C34)}$$

This expression for the change in $\langle X_1\rangle$ is exact, but it depends on the value actually measured in the nth measurement. More useful for discussing the behavior of X_1 would be some sort of expected value for the change in $\langle X_1\rangle$. Indeed, throughout the rest of this Appendix, we shall want to deal with such expected changes between measurements and with expected changes over an entire sequence of measurements.

Defining such expected changes requires introducing a new type of average, which we shall denote by a superposed bar. A superposed bar was used previously in Eq. (C24a) to denote the mean value of ξ_n. There it meant an average over an ensemble of identical oscillators which began the nth measurement in the same state; such an average is, of course, equivalent to an expectation value. In all *other* applications throughout the rest of this Appendix, a superposed bar will denote an average over an ensemble of identical oscillators which begin in the same state before the *initial* measurement in a sequence; this "barred average" is a generalization of the usual notion of expectation value.

One must keep in mind that the mean value $\bar{\xi}_n$ is *not* an average over this second type of ensemble; rather, each oscillator in the ensemble has its own value of $\bar{\xi}_n$—a value which depends on the results of previous measurements for that particular oscillator [see Eq. (C40)]. On the other hand, all oscillators in the ensemble do have the same set of values for the uncertainties $(\Delta X_1)_n$ [Eq. (C33a)] and the measurement errors $\Delta\xi_n$ [Eq. (C24b)]. This makes it easy to apply the average to the differences $(\xi_n-\bar{\xi}_n)$; these differences are statistically independent quantities with mean zero and with correlation matrix

$$\overline{(\xi_k-\bar{\xi}_k)(\xi_l-\bar{\xi}_l)}=(\Delta\xi_k)^2\delta_{kl}. \quad \text{(C35)}$$

We can now return to Eq. (C34) and apply the concept of a barred average. We first note that the mean change of $\langle X_1\rangle$ in a given measurement is zero ($\overline{\langle X_1\rangle_n-\langle X_1\rangle_{n-1}}=0$). However, the change does have an rms value, which can be thought of as the expected magnitude of the change in $\langle X_1\rangle$:

$$(\delta X_1)_n\equiv[\overline{(\langle X_1\rangle_n-\langle X_1\rangle_{n-1})^2}]^{1/2}=(\Delta X_1)_{n-1}[1+\sigma^2/(\Delta X_1)^2_{n-1}]^{-1/2} \quad \text{(C36)}$$

[cf. Eq. (C24b)]. Note that $(\delta X_1)_n\leq(\Delta X_1)_{n-1}$—i.e., the expected change in $\langle X_1\rangle$ is always less than or equal to the variance of $\hat{X}_1$ at the beginning of the measurement. If $(\Delta X_1)_{n-1}\gtrsim\sigma$, then $(\delta X_1)_n\sim(\Delta X_1)_{n-1}$; however, if $(\Delta X_1)_{n-1}\ll\sigma$, then $(\delta X_1)_n\simeq(\Delta X_1)^2_{n-1}/\sigma\ll(\Delta X_1)_{n-1}$.

To make further progress, we must specify the oscillator state $|\chi_0\rangle$ before the initial ($n=1$) measurement in the sequence. We choose a state of the form of Eq. (C31): a minimum-uncertainty state with $\langle X_2\rangle_0=0$ ($a_0=\mu_0=0$), with $(\Delta X_1)_0\gg\sigma$, and with $\langle X_1\rangle_0$ arbitrary. A good example of such a state is a coherent state [$(\Delta X_1)_0=(\Delta X_2)_0=(\hbar/2m\omega)^{1/2}\gg\sigma$ if $\eta\gg1$]. The oscillator can be prepared in a coherent state using high-precision "amplitude-and-phase" techniques (see discussion in Sec. II. B). Throughout the following we neglect terms of order $\sigma/(\Delta X_1)_0$.

The first measurement in the sequence serves essentially as a "state-preparation measurement." Its result is highly uncertain, but it leaves the oscillator in a state with $\langle X_1\rangle_1=\xi_1$ and $(\Delta X_1)_1=\sigma$ [Eqs. (C33)]. We assume there is no feedback during the first measurement ($\alpha_0=0$).

Subsequent measurements are the ones of real interest. Equations (C33) can be iterated to obtain the constants describing the oscillator state after the nth measurement:

$$(\Delta X_1)_n=\sigma/\sqrt{n}, \quad \text{(C37a)}$$

$$\langle X_1\rangle_n=\frac{1}{n}\sum_{k=1}^{n}\xi_k, \quad \text{(C37b)}$$

$$a_n=\eta^2\sum_{k=1}^{n}(3\xi_k-\tfrac{4}{3}\alpha_{k-1}), \quad \text{(C37c)}$$

$$\mu_n=\tfrac{5}{6}\eta^2 n \quad \text{(C37d)}$$

($n \geq 1$); Eq. (C24b), together with (C37a), gives the likely error of the nth measurement:

$$\Delta\xi_n = \sigma[n/(n-1)]^{1/2}, \quad n \geq 2; \tag{C38}$$

and Eqs. (C37) applied to Eqs. (C32) give the expectation value and variance of $\hat{X}_2$ after the nth measurement:

$$\langle X_2\rangle_n = \tfrac{4}{3}\eta^2 \sum_{k=1}^{n} (\xi_k - \alpha_{k-1}), \tag{C39a}$$

$$(\Delta X_2)_n = \tfrac{1}{3}\sqrt{34n}\,\eta(\hbar/2m\omega)^{1/2} \quad (n \geq 1). \tag{C39b}$$

6. Discussion of results

Equations (C37)–(C39) provide a complete description of the sequence of measurements; our task now is to discuss their implications.

We first note that the variances of $\hat{X}_1$ and $\hat{X}_2$ change in a completely deterministic way, independent of the actual measured values. On the other hand, the changes in the expectation values are entirely dependent on the measured values. Indeed, the expectation value of $\hat{X}_1$ after a given measurement is simply the arithmetic mean of all previously measured values. This last statement means that the experimenter knows in advance the expected result of each measurement after the first—i.e.,

$$\bar{\xi}_n = \langle X_1\rangle_{n-1} = \frac{1}{n-1}\sum_{k=1}^{n-1} \xi_k. \tag{C40}$$

This is the finite-coupling analog of the situation analyzed in Sec. IV.A. There we assumed infinite coupling, and the experimenter could predict *exactly* the result of each measurement after the first. Here we have finite coupling, the experimenter knows the *expected* result of each measurement, but the actual result is likely to differ from the expected by an amount $\Delta\xi_n \simeq \sigma$. Equation (C40) also describes the situation one wants for measuring a classical force, because one detects the force by looking at the difference between the actual measured value and the (known) expected result (cf. Sec. III.B).

Given a set of measured values, one can calculate the changes in the expectation value of $\hat{X}_1$ using Eq. (C37b). Exact this may be, but enlightening it is not. To gain insight we look at expected changes in $\langle X_1\rangle$, and to do that we begin by writing Eq. (C34) in the form

$$\langle X_1\rangle_n - \langle X_1\rangle_{n-1} = (\xi_n - \bar{\xi}_n)/n. \tag{C41}$$

The expected change in $\langle X_1\rangle$ is

$$(\delta X_1)_n = \Delta\xi_n/n = \sigma/[n(n-1)]^{1/2}. \tag{C42}$$

The expectation value of $\langle \hat{X}_1\rangle$ "jumps" at each measurement. The "jumps" add randomly, but their expected size decreases so rapidly that after many measurements $\langle X_1\rangle$ is likely to have wandered only a distance $(\Delta X_1)_1 = \sigma$ from its value after the first measurement—i.e.,

$$[\overline{(\langle X_1\rangle_n - \langle X_1\rangle_1)^2}]^{1/2} = \left(\sum_{k=2}^{n} \frac{(\Delta\xi_k)^2}{k^2}\right)^{1/2} = \sigma\left(1 - \frac{1}{n}\right)^{1/2} \tag{C43}$$

[cf. Eqs. (C41), (C35), and (C38)]. This means that the results of all measurements after the first cannot determine $\langle X_1\rangle_1$ more accurately than $(\Delta X_1)_1$. [One can easily show that the jumps prevent measurements after the nth from determining $\langle X_1\rangle_n$ more accurately than $(\Delta X_1)_n$.]

This behavior of X_1 can be summarized as follows: After the first measurement the oscillator is in a state with $\langle X_1\rangle_1 = \xi_1$ and $(\Delta X_1)_1 = \sigma$. For the next few measurements $\langle X_1\rangle$ jumps around within a region $\xi_1 \pm$(a few) $\times\sigma$, while ΔX_1 gets smaller at each measurement. As the sequence proceeds the jumps of $\langle X_1\rangle$ become smaller and smaller, $\langle X_1\rangle$ "zeros in" on some particular value $\langle X_1\rangle_\infty$, ΔX_1 goes to zero, and the probability distribution $|\chi(\xi)|^2$ approaches a delta function at $\langle X_1\rangle_\infty$. The final value $\langle X_1\rangle_\infty$ is likely to be within the region $\xi_1 \pm$(a few) $\times\sigma$.

We now turn to the behavior of X_2, and we begin by noting that one can associate with X_2 a characteristic "quantum step size" $\sim \sigma^{-1}(\hbar/2m\omega) = \eta(\hbar/2m\omega)^{1/2}$, obtained from the basic accuracy σ of X_1 measurement and the uncertainty principle, Eq. (2.9a).

The expectation value of $\hat{X}_2$ changes at each measurement, and the change is given by

$$\langle X_2\rangle_n - \langle X_2\rangle_{n-1} = \tfrac{4}{3}\eta^2(\xi_n - \alpha_{n-1}). \tag{C44}$$

These "kicks" to X_2 are essentially classical. Indeed, Eq. (C44) is precisely the classical displacement of X_2, which our measurement system would produce in a classical oscillator with $X_1 = \xi_n$, during the time interval between t_{n-1} and $t_n = t_{n-1} + \tau$; cf. Eqs. (C1) viewed classically, together with $Q(t_{n-1}) = \Pi(t_{n-1}) = 0$, and Eq. (C26). In the absence of feedback, the kicks (C44) accumulate and $\langle X_2\rangle$ runs away. However, feedback can eliminate this "classical runaway of X_2," because the measured value ξ_n of X_1 tells one precisely the kick given $\langle X_2\rangle$ during the nth measurement. The simplest feedback is to let $\alpha_n = \xi_n$; then the feedback between t_n and t_{n+1} cancels the kick given X_2 in the nth measurement.

One can do much better by choosing the feedback so that at each measurement it not only cancels the previous kick but also attempts to cancel the current kick. The feedback cannot cancel the current kick precisely, because to do so would require knowing the result of the measurement. However, one can try to guess the result, and the best guess is the expected result (C40). The resulting feedback has the form

$$\alpha_1 = 2\xi_1,$$

$$\alpha_n = \xi_n - \frac{1}{n-1}\sum_{k=1}^{n-1}\xi_k + \frac{1}{n}\sum_{k=1}^{n}\xi_k, \quad n > 1. \tag{C45}$$

With this feedback the expectation value of $\hat{X}_2$ after n measurements is

$$\langle X_2\rangle_n = \tfrac{4}{3}\eta^2(\xi_n - \bar{\xi}_n), \tag{C46}$$

a displacement with mean zero and with rms value

$$[\overline{\langle X_2\rangle_n^2}]^{1/2} = \tfrac{4}{3}\eta^2\Delta\xi_n = \tfrac{4}{3}[n/(n-1)]^{1/2}\eta(\hbar/2m\omega)^{1/2}, \quad n \geq 2. \tag{C47}$$

The effectiveness of the feedback is evident from its

ability to keep $\langle X_2\rangle$ within one "quantum step" of zero.

Effective though the feedback may be, it cannot prevent the huge, unpredictable, quantum-mechanical kicks given X_2 by precise measurements of X_1. As Eq. (C39b) shows, the effect of these kicks appears in the variance of $\hat{X}_2$, which grows as $\sqrt{n}$—behavior which suggests that of a classical random-walk variable. The step size is $\frac{1}{3}\sqrt{34}\,\eta(\hbar/2m\omega)^{1/2}$, in agreement with what one predicts from the uncertainty principle. This "random walk of X_2" means that the energy in the oscillator grows as the sequence proceeds:

$$\langle \hat{H}_0\rangle_n \simeq \tfrac{1}{2}m\omega^2(\Delta X_2)^2_n = \tfrac{17}{18}\eta^2 n\hbar\omega \qquad \text{(C48)}$$

[Eqs. (2.2) and (2.6)]. The source of the energy is the generator. Interaction with the generator can add energy to or remove energy from the oscillator, but on the average energy is added. Practical implications of the random walk of X_2 are considered in Paper II.

The analysis in this Appendix has emphasized the possibility of making quick measurements of X_1, but nothing restricts the analysis to this case. It applies equally well to measurements of X_1 which, because of weak coupling, require a long time to achieve good accuracy. This point is made clear by introducing a new constant

$$\varepsilon^2 \equiv \eta^2(\omega\tau)^{-3} = K^2/8mL\omega^4\,, \qquad \text{(C49)}$$

which is a dimensionless measure of the coupling strength. Written in terms of ε, the fundamental accuracy becomes

$$\sigma = \varepsilon^{-1}(\omega\tau)^{-3/2}(\hbar/2m\omega)^{1/2}\,. \qquad \text{(C50)}$$

If $\varepsilon \gg 1$, a measurement much shorter than a period can beat the standard quantum limit, Eq. (2.16)[13]; but if $\varepsilon \ll 1$, beating the standard limit requires a measurement many periods long. Regardless of how small ε may be, the basic accuracy (C50) can be made as small as one desires (in principle) by choosing τ large enough. Long measurement times yield arbitrarily good accuracy because $\hat{X}_1$ is completely isolated from noise in the measuring apparatus (cf. Sec. IV.B). The constant ε plays an important role in Appendix D, and it and its relatives are considered extensively in Paper II.

7. Analysis of imprecise readout systems

One possible objection to the above analysis is its treatment of the readout system. We have assumed that when the readout determines a value for the meter coordinate, it localizes the meter coordinate precisely at the measured value. Of course, no real readout system can achieve such arbitrarily good precision. One way to handle this difficulty is to do a better job of analyzing the readout: Specify in detail the design of a realistic readout system, and include all or part of the readout in the exact quantum-mechanical analysis. The resulting analysis is likely to be difficult, if not impossible.

Fortunately, there is an easier and more general approach. In this approach the imprecision of the readout system is described by a (classical) conditional probability distribution $W(Q\,|\,Q_n)$. The distribution $W(Q\,|\,Q_n)$ can be thought of as giving the probability $W(Q\,|\,Q_n)dQ$ that, when the readout determines a value Q_n for the meter coordinate, the meter is actually located between Q and $Q+dQ$.

The introduction of $W(Q\,|\,Q_n)$ can be justified by considering a simple model for the readout system. The first three-quarters of this section [through Eq. (C59)] will present that model and will show how it gives rise to $W(Q\,|\,Q_n)$. The last quarter will assume a simple form for $W(Q\,|\,Q_n)$, and from it will derive results for the measurement errors and variances in a sequence of measurements with an imprecise readout system.

In our simple model for the readout system, the first stage is a "readout meter": a one-dimensional, quantum-mechanical "free mass" with generalized coordinate $\hat{\mathcal{Q}}$, generalized momentum $\hat{\mathcal{P}}$, and generalized mass M. The readout meter is coupled to the meter by coordinate-coordinate coupling; hence, the total Hamiltonian for the oscillator, the meter, and the readout meter is

$$\hat{H}_T = \hat{H}[\text{Eqs. (C1)}] + \mathcal{K}\hat{Q}\,\hat{\mathcal{Q}} + \hat{\mathcal{P}}^2/2M\,, \qquad \text{(C51)}$$

where $\mathcal{K}$ is a coupling constant. We shall include the readout meter in the quantum-mechanical analysis. The readout meter is coupled to subsequent stages of the readout system in such a way that, at designated moments of time, the subsequent stages can "read out" a value for the readout-meter coordinate. We shall idealize these readouts of $\mathcal{Q}$ as arbitrarily precise, essentially instantaneous measurements. Then we need not treat the subsequent stages of the readout system quantum mechanically—i.e., we can place the quantum-classical cut of our analysis between the readout meter and the subsequent stages of the readout system.

The scenario envisioned for the nth measurement divides neatly into two parts. During the first part, lasting from t_{n-1} to $t_{n-1}+\tau$, the oscillator and meter interact via the interaction Hamiltonian $K(\hat{X}_1 - \alpha_{n-1})\hat{Q}$ [Eq. (C1d)] just as in the previous analysis. During the second part, lasting from $t_{n-1}+\tau$ to $t_n = t_{n-1}+\tau+\bar{\tau}$ (note that t_n is defined differently than in the previous analysis), the meter and the readout meter interact via the interaction Hamiltonian $\mathcal{K}\hat{Q}\,\hat{\mathcal{Q}}$ [Eq. (C51)]. (The coupling "constants" have the following functional form for the nth measurement: $K = \mathcal{K} = 0$ for $t < t_{n-1}$ and $t > t_n$; $K = \text{const} \neq 0$, $\mathcal{K} = 0$ for $t_{n-1} < t < t_{n-1}+\tau$; $K = 0$, $\mathcal{K} = \text{const} \neq 0$ for $t_{n-1}+\tau < t < t_n$.) At time t_n the subsequent stages of the readout system read out a value of $\mathcal{Q}$, from which the experimenter infers a value of Q (and X_1). The three operations of (i) interaction between meter and readout meter, (ii) readout of $\mathcal{Q}$, and (iii) inference of a value for Q, together constitute what was called the "readout of the meter coordinate" in the previous analysis. After the nth measurement the meter is thrown away; a new meter is used for the next measurement.

The discarding of the meter at the end of each measurement is an important feature of our analysis. Unless we keep track of the states of the meters discarded

[13] For quick measurements ($\tau \lesssim \omega^{-1}$) it is more reasonable to compare the accuracy σ to $(\hbar\tau/m)^{1/2}$, the standard quantum limit for measurements of free-mass position [Eq. (3.2)]. Beating this standard limit requires even stronger coupling than is required to beat the standard quantum limit for amplitude-and-phase measurements.

in previous measurements, it will turn out that the entire system cannot be described by a pure state; instead it must be described by a mixed state. Thus the analysis is most conveniently carried out using density operators. During the nth measurement the state of the total system—oscillator, meter, and readout meter—is specified by a density operator $\hat{\rho}_T(t)$ with associated density matrix

$$\rho_T(\xi,\xi';Q,Q';\mathcal{Q},\mathcal{Q}';t)\equiv\langle\xi,Q,\mathcal{Q};t\,|\,\hat{\rho}_T(t)\,|\,\xi',Q',\mathcal{Q}';t\rangle,$$

where the states $|\xi,Q,\mathcal{Q};t\rangle$ are the obvious generalization of the states $|\xi,Q;t\rangle$. The density matrix has the interpretation that $\rho_T(\xi,\xi;Q,Q;\mathcal{Q},\mathcal{Q};t)\,d\xi\,dQ\,d\mathcal{Q}$ is the probability at time t of simultaneously finding the readout meter between $\mathcal{Q}$ and $\mathcal{Q}+d\mathcal{Q}$, the meter between Q and $Q+dQ$, and the oscillator with X_1 between ξ and $\xi+d\xi$. The total density operator evolves according to

$$\hat{\rho}_T(t)=\hat{U}_T(t,t_0)\hat{\rho}_T(t_0)\hat{U}_T^\dagger(t,t_0),$$

where $\hat{U}_T(t,t_0)$ is the time-development operator for the total Hamiltonian (C51).

During the first part of the nth measurement ($t_{n-1}<t<t_{n-1}+\tau$), we need only be concerned with the state of the oscillator and meter. Their state is specified by a density operator $\hat{\rho}_{n-1}(t)$, which evolves according to

$$\hat{\rho}_{n-1}(t)=\hat{U}(t,t_{n-1})\hat{\rho}_{n-1}(t_{n-1})\hat{U}^\dagger(t,t_{n-1})$$

[cf. Eq. (C13)], and which has density matrix

$$\rho_{n-1}(\xi,\xi';Q,Q';t)\equiv\langle\xi,Q;t\,|\,\hat{\rho}_{n-1}(t)\,|\,\xi',Q';t\rangle.$$

We now analyze the components of the nth measurement in greater detail. The oscillator begins the measurement (at time t_{n-1}) in a state with density matrix $\Upsilon_{n-1}(\xi,\xi')$, and the meter is prepared in the (pure) Gaussian state (C18). The initial density operator $\hat{\rho}_{n-1}(t_{n-1})$ has density matrix

$$\rho_{n-1}(\xi,\xi';Q,Q';t_{n-1})=\Upsilon_{n-1}(\xi,\xi')\Phi(Q)\Phi^*(Q').$$

The oscillator and meter interact as in the previous analysis for a time τ; the evolution of ρ_{n-1} during this time can be inferred from the evolution of the corresponding wave function [Eqs. (C16) and (C19)].

At time $t_{n-1}+\tau$ the readout meter is prepared in a (pure) state with wave function $\Theta(\mathcal{Q})$, which has $\langle\hat{\mathcal{Q}}\rangle=\langle\hat{\mathcal{P}}\rangle=0$. For the moment we leave the precise form of $\Theta(\mathcal{Q})$ unspecified. The total density matrix at time $t_{n-1}+\tau$ is

$$\rho_T(\xi,\xi';Q,Q';\mathcal{Q},\mathcal{Q}';t_{n-1}+\tau)=\rho_{n-1}(\xi,\xi';Q,Q';t_{n-1}+\tau)\Theta(\mathcal{Q})\Theta^*(\mathcal{Q}').$$

The expectation value and variance of $\hat{Q}$ at time $t_{n-1}+\tau$ are denoted $\langle\hat{Q}\rangle_i$ and $(\Delta Q)_i$; they are given by Eqs. (C20). During the subsequent interval of duration $\tilde{\tau}$ ($t_{n-1}+\tau<t<t_n=t_{n-1}+\tau+\tilde{\tau}$), the meter and the readout meter interact; the total Hamiltonian is given by Eq. (C51) with $K=0$. We make three assumptions about the evolution of the system during this time:

Assumption i: $\tilde{\tau}\ll\tau$;

Assumption ii: $\dfrac{\mathcal{K}^2\tilde{\tau}^4}{LM}\ll 1\left[\left(\dfrac{\hbar M}{\mathcal{K}^2\tilde{\tau}^3}\right)^{1/2}\gg\left(\dfrac{\hbar\tilde{\tau}}{L}\right)^{1/2}\right]$;

Assumption iii: $\left(\dfrac{\hbar\tilde{\tau}}{M}\right)^{1/2}\lesssim(\Delta\mathcal{Q})_i\ll\left(\dfrac{\hbar\tilde{\tau}}{M}\right)^{1/2}\left(\dfrac{LM}{\mathcal{K}^2\tilde{\tau}^4}\dfrac{\tau}{\tilde{\tau}}\right)^{1/2}$,

$$(\Delta\mathcal{P})_i\ll\left(\frac{\hbar M}{\tilde{\tau}}\right)^{1/2}\left(\frac{LM}{\mathcal{K}^2\tilde{\tau}^4}\frac{\tau}{\tilde{\tau}}\right)^{1/2}.$$

Here the subscript "i" denotes the value at $t_{n-1}+\tau$. These assumptions guarantee that the meter coordinate remains essentially undisturbed by the evolution of the entire system during the time $\tilde{\tau}$. Assumption i guarantees that the meter does not evolve significantly under the influence of its own Hamiltonian. Assumptions ii and iii guarantee that the "back action" of the readout meter onto the meter coordinate is negligible.

Assumptions ii and iii can be viewed in another way. They imply that the readout meter does not do a very good job of measuring the meter coordinate—i.e., the readout meter is far from being a "quantum-limited measuring device." Assumption ii guarantees that, in measuring Q, the best accuracy the readout meter can achieve is far worse than the standard quantum limit $(\hbar\tilde{\tau}/L)^{1/2}$ [cf. Eq. (3.2)]. Assumption iii allows $(\Delta\mathcal{Q})_i$ and $(\Delta\mathcal{P})_i$ to be much greater than the optimum uncertainties for a measurement of duration $\tilde{\tau}$. Thus we do not place stringent demands on the performance of the readout meter. That this is intimately connected with the absence of back action onto Q should not be surprising.

With assumptions i–iii, the evolution of the total system is precisely analogous to the evolution of the meter coupled to $\hat{X}_1$ of the oscillator [Eqs. (C16)]. The total density matrix is given by

$$\begin{aligned}\rho_T(\xi,\xi';Q,Q';\mathcal{Q},\mathcal{Q}';t)&=\rho_{n-1}(\xi,\xi';Q,Q';t_{n-1}+\tau)\\&\quad\times\exp\left(-\frac{i}{\hbar}\left[f(Q,\mathcal{Q},t)-f(Q',\mathcal{Q}',t)\right]\right)\\&\quad\times\Theta_{\text{free}}\left(\mathcal{Q}+\frac{\mathcal{K}v^2}{2M}Q,t\right)\Theta^*_{\text{free}}\left(\mathcal{Q}'+\frac{\mathcal{K}v^2}{2M}Q',t\right),\end{aligned}\tag{C52a}$$

$$f(Q,\mathcal{Q},t)\equiv(\mathcal{K}^2v^3/6M)Q^2+\mathcal{K}vQ\mathcal{Q},\tag{C52b}$$

where $v\equiv t-t_{n-1}-\tau$, and where $\Theta_{\text{free}}(\mathcal{Q},t)$ gives the evolution of a free readout meter with initial state $\Theta(\mathcal{Q})$ [analog of Eqs. (C16b) and (C16c)].

During the interval $\tilde{\tau}$ the readout meter "swings" due to its interaction with the meter. At time t_n the expectation value and variance of $\hat{\mathcal{Q}}$ become

$$\langle\hat{\mathcal{Q}}\rangle(t_n)=-(\mathcal{K}\tilde{\tau}^2/2M)\langle\hat{Q}\rangle_i,\tag{C53a}$$

$$\Delta\mathcal{Q}(t_n)=[(\Delta\mathcal{Q})^2_{\text{free}}(t_n)+(\mathcal{K}\tilde{\tau}^2/2M)^2(\Delta Q)_i^2]^{1/2},\tag{C53b}$$

where $(\Delta\mathcal{Q})_{\text{free}}(t_n)$ is the variance of a free readout meter.

At time t_n the subsequent stages of the readout system read out a value $\mathcal{Q}_n$ for the readout-meter coordinate. Using Eq. (C53a) the experimenter infers a value

$$Q_n=-(2M/\mathcal{K}\tilde{\tau}^2)\mathcal{Q}_n\tag{C54}$$

for the meter coordinate. In the terminology of the previous analysis, Q_n is the result of the "readout of the meter coordinate." From Q_n the experimenter infers a value ξ_n for X_1 just as before [Eq. (C21)]. The probability distribution of ξ_n (referred to an ensemble

of identical systems which begin the nth measurement in the same state) is easily obtained from the probability distribution of $\mathcal{Q}$:

$$P(\xi_n)=\frac{\mathcal{K}K\bar{\tau}^2\tau^2}{4ML}\int d\xi\, dQ\, \rho_T(\xi,\xi;Q,Q;\mathcal{Q}_n,\mathcal{Q}_n;t_n)$$
$$=(2\pi\sigma^2)^{-1/2}\int d\xi\, dQ\, \Upsilon_{n-1}(\xi,\xi)$$
$$\times\exp\left(-\frac{[\xi-\alpha_{n-1}+(2L/K\tau^2)Q]^2}{2\sigma^2}\right)W(Q\,|\,Q_n) \quad \text{(C55)}$$

[cf. Eq. (C22)], where the conditional probability distribution $W(Q\,|\,Q_n)$ is defined by

$$W(Q\,|\,Q_n)\equiv\frac{\mathcal{K}\bar{\tau}^2}{2M}\left|\,\Theta_{\text{free}}\left(\frac{\mathcal{K}\bar{\tau}^2}{2M}(Q-Q_n),t_n\right)\right|^2. \quad \text{(C56)}$$

$W(Q\,|\,Q_n)$ has mean Q_n and variance $\sigma_W=(2M/\mathcal{K}\bar{\tau}^2)(\Delta\mathcal{Q})_{\text{free}}(t_n)$.

The mean and variance of ξ_n (averages over the ensemble) are

$$\bar{\xi}_n=\langle X_1\rangle_{n-1}, \quad \text{(C57a)}$$
$$\Delta\xi_n=[(2L/K\tau^2)^2\sigma_W^2+\sigma^2+(\Delta X_1)^2_{n-1}]^{1/2} \quad \text{(C57b)}$$

[cf. Eqs. (C24)]. The measurement error $\Delta\xi_n$ is the same as in the previous analysis, except that it is augmented by a term which accounts for the imprecision of the readout meter. Even if the readout meter is extremely imprecise [$\sigma_W\gg(\hbar\tau/L)^{1/2}$], it is still true that the measurement of X_1 can be arbitrarily accurate when K is made arbitrarily large.

When the subsequent stages of the readout system read out the value $\mathcal{Q}_n$, they localize the readout meter precisely at $\mathcal{Q}_n$. This "reduction of the wave function" means that immediately after the readout the density matrix of the oscillator-meter system is

$$\rho_a(\xi,\xi';Q,Q';t_n)=\mathcal{C}\rho_T(\xi,\xi';Q,Q';\mathcal{Q}_n,\mathcal{Q}_n;t_n), \quad \text{(C58)}$$

where $\mathcal{C}$ is a normalization constant. After the readout we throw the meter away, and we prepare a new meter for use in the next measurement. Throughout all subsequent measurements in the sequence, we shall not be interested in computing any expectation values which involve observables of the discarded meter. To compute any other expectation value we must "take the trace" of the density matrix on Q. Therefore, insofar as any future expectation values of interest are concerned, we can take the trace on Q now—i.e., we can replace the density matrix (C58) with a density matrix that describes only the oscillator:

$$\Upsilon_n(\xi,\xi')=\int\rho_a(\xi,\xi';Q,Q;t_n)\,dQ$$
$$=\mathcal{C}\int W(Q\,|\,Q_n)\rho_{n-1}(\xi,\xi';Q,Q;t_{n-1}+\tau)\,dQ, \quad \text{(C59)}$$

where $\mathcal{C}$ is another normalization constant. Equation (C59) gives the initial oscillator state for the $(n+1)$th measurement [cf. Eq. (C29)].

The key results of our analysis of an imprecise readout system are Eqs. (C55) and (C59). They justify our claim that the imprecision of the readout can be described by a classical probability distribution: Eq. (C55) shows how the readout imprecision contributes to the measurement error, and Eq. (C59) shows how the readout imprecision "smears out" the "reduction of the wave function." In the limit that the readout meter is arbitrarily precise [$W(Q\,|\,Q_n)=\delta(Q-Q_n)$], Eqs. (C55) and (C59) reduce to the corresponding equations of the previous analysis [cf. Eqs. (C22) and (C29)]. Indeed, this analysis justifies our previous treatment of an arbitrarily precise readout system—i.e., it justifies the procedure of "reducing the wave function" after each arbitrarily precise readout.

Two features of this analysis deserve special emphasis. The first is that we have made assumptions which guarantee that the meter coordinate is essentially undisturbed by the interaction with the readout meter. Formally, this means that the total density matrix (C52) splits cleanly into a product of two terms: (i) a density matrix for the free oscillator-meter system; and (ii) a function which depends only on the meter coordinate and the readout meter coordinate. The second feature is that we throw away the meter after each measurement. Both these features are necessary for defining $W(Q\,|\,Q_n)$; and it is the loss of information that occurs when the meter is discarded which allows us to identify the oscillator state after the measurement, and which converts an initial, pure oscillator state into a mixed state.

We must specify a particular form for $W(Q\,|\,Q_n)$ in order to use Eq. (C59) to analyze a sequence of measurements. A reasonable form is a Gaussian with mean Q_n and variance $\sigma_W=\gamma(1-\gamma^2)^{-1/2}(\hbar\tau/L)^{1/2}$ $(0\leq\gamma<1)$. This is the form $W(Q\,|\,Q_n)$ would have if $\Theta(\mathcal{Q})$ were a Gaussian wave packet. When $\gamma^2\simeq0.5$ the readout imprecision contributes about the same amount to the measurement error as the meter uncertainty [cf. Eq. (C57b)]. Using this Gaussian for $W(Q\,|\,Q_n)$, we have integrated Eq. (C59) to obtain

$$\Upsilon_n(\xi,\xi')=\mathcal{B}\Upsilon_{n-1}(\xi,\xi')\exp\left\{-\frac{1}{4\sigma^2}\,[(1+4\gamma^2)(\xi^2+\xi'^2)-10\gamma^2\xi\xi'-2(1-\gamma^2)\xi_n(\xi+\xi')]\right.$$
$$\left.+(i/\hbar)m\omega\eta^2[3\xi_n(1-\gamma^2)(\xi-\xi')-\tfrac{4}{3}\,\alpha_{n-1}(\xi-\xi')-\tfrac{5}{6}(1-\tfrac{9}{5}\gamma^2)(\xi^2-\xi'^2)]\right\}, \quad \text{(C60)}$$

where $\mathcal{B}$ is a normalization constant. This equation is a generalization of the fundamental equation (C30); it simplifies to (C30) when $\gamma=0$.

Equation (C60) can be used to analyze a sequence of measurements. In particular, it can be used to analyze a sequence in which the oscillator begins in the same (pure) state as in the previous analysis. The results for the expectation values $\langle X_1\rangle_n$ and $\langle X_2\rangle_n$ are the same as before [Eqs. (C37b) and (C39a)]; but Eq. (C38) for the measurement error becomes

$$\Delta\xi_n=[\sigma/(1-\gamma^2)^{1/2}]\,[n/(n-1)]^{1/2} \text{ for } n\geq2, \quad \text{(C61)}$$

and Eqs. (C37a) and (C39b), which give the evolution of ΔX_1 and ΔX_2, are changed to

$$(\Delta X_1)_n = \sigma/[n(1-\gamma^2)]^{1/2}, \tag{C62a}$$

$$(\Delta X_2)_n = \tfrac{1}{3}(34n)^{1/2}\eta(\hbar/2m\omega)^{1/2}[(1-\tfrac{9}{17}\gamma^2)/(1-\gamma^2)]^{1/2} \tag{C62b}$$

($n \geq 1$). For reasonable values of γ ($\gamma^2 \simeq 0.5$), the decrease of ΔX_1 and the growth of ΔX_2 are not markedly different from the results of the previous analysis.

APPENDIX D: SINGLE-TRANSDUCER BACK-ACTION-EVADING MEASUREMENTS OF X_1: A FULLY QUANTUM-MECHANICAL ANALYSIS[14]

1. Introduction

In this Appendix we give a fully quantum-mechanical analysis of a single-transducer back-action-evading measurement of the X_1 of a harmonic oscillator (see Sec. II.F.3). We consider a simplified version of a real measuring apparatus, analyze the measurement process quantum mechanically, and thereby demonstrate that in principle such single-transducer measurements can beat the standard quantum limit $\Delta X_1 = (\hbar/2m\omega)^{1/2}$ [Eq. (2.16)].

Single-transducer back-action-evading measurements are considered extensively in Paper II, where they are analyzed using semiclassical techniques. Those semiclassical analyses are to be preferred in almost every way over the analysis given here: They are more realistic and more adaptable, and they provide more detailed information. However, the reader might harbor lingering doubts about the validity of applying semiclassical techniques to measurements which purport to beat the standard quantum limit. The purpose of this Appendix is to remove such doubts by analyzing quantum mechanically a simple example of a single-transducer back-action-evading measurement.

The analysis we give here is similar to the analysis in Sec. III.B and Appendix C. In particular, the measuring apparatus is the same. It consists of a generator, which provides the time dependence in the interaction Hamiltonian; a meter, which is a one-dimensional quantum-mechanical "free mass" coupled to the oscillator by the generator; and a readout system, which reads out the position of the meter. Only the meter will be treated quantum mechanically.

The difference between here and Sec. III.B lies in the way the meter is coupled to the oscillator. In Sec. III.B the meter was perfectly coupled to $\hat{X}_1$; here the meter is coupled to $\hat{X}_1$ only in a time-averaged sense. The total Hamiltonian for the oscillator coupled to the meter via the classical generator is given by Eqs. (3.16), except that in the interaction Hamiltonian the momentum coupling is omitted:

$$\hat{H}_I = K\hat{Q}\hat{x}\cos\omega t = \tfrac{1}{2}K\hat{Q}[\hat{X}_1(1+\cos 2\omega t) + \hat{X}_2 \sin 2\omega t] \tag{D1}$$

[cf. Eq. (2.42a)]. Systems which in principle are governed by Hamiltonian (3.16) are considered in Appendix B. They can be modified easily to have the Hamiltonian considered here; essentially, the modification consists of deleting the momentum transducer.

The motivation for considering single-transducer back-action-evading measurements is the problem of weak coupling. In Sec. III.B and in Appendix C we showed that back-action-evading measurements with perfect $\hat{X}_1$ coupling can achieve arbitrarily good accuracy in an arbitrarily short time. However, such quick measurements (short compared to an oscillator period) require that the measuring apparatus be strongly coupled to the oscillator. In Appendix C we introduced a constant

$$\varepsilon^2 \equiv K^2/8mL\omega^4 \tag{D2}$$

[cf. Eq. (C49)], which provides a dimensionless measure of coupling strength for a simple "free-mass" meter coupled to an oscillator. Quick measurements require $\varepsilon \gg 1$. If $\varepsilon \lesssim 1$, beating the standard quantum limit of Eq. (2.16) requires a measurement time longer than a period.

In real experiments it is often quite difficult to achieve strong coupling. If one is stuck with weak coupling ($\varepsilon \ll 1$), then the required long measurement time allows one to avoid coupling perfectly to $\hat{X}_1$ and permits one instead to couple to $\hat{X}_1$ in a time-averaged sense. In particular, one can omit one of the two transducers (position or momentum) required for perfect coupling, with a consequent simplification in the design and construction of the measuring apparatus. One modulates the output of the remaining transducer so that at some frequency the modulated output carries the desired information about X_1 with very little contamination from X_2, and one then runs the modulated output through a filter which picks out the desired frequency. (See Paper II for details; and see Thorne *et al.*, 1979 for a semirealistic example.) The Hamiltonian (3.16a)–(3.16c), with interaction term (D1), is the simplest example of this procedure: The momentum transducer is omitted, the modulation of the position transducer is a sinusoid at the oscillator frequency, the desired X_1 signal is at zero frequency, and the meter—a zero-frequency harmonic oscillator—serves as a filter at zero frequency.

Since single-transducer measurements are useful only in the case of weak coupling, we assume $\varepsilon \ll 1$ throughout this Appendix.

2. The analysis

The analysis proceeds by solving for the evolution of the appropriate operators in the Heisenberg picture. The Hamiltonian (3.16a)–(3.16c), with interaction term (D1), yields the following Heisenberg equations of motion:

$$d\hat{X}_1/dt = (K/2m\omega)\hat{Q}\sin 2\omega t, \tag{D3a}$$

$$d\hat{X}_2/dt = -(K/2m\omega)\hat{Q}(1+\cos 2\omega t), \tag{D3b}$$

$$d\hat{Q}/dt = \hat{\Pi}/L, \tag{D3c}$$

$$d\hat{\Pi}/dt = -\tfrac{1}{2}K[\hat{X}_1(1+\cos 2\omega t) + \hat{X}_2 \sin 2\omega t]. \tag{D3d}$$

The crucial difference between these equations and those for perfect $\hat{X}_1$ coupling [Eqs. (3.17)] is that $\hat{X}_1$ is not completely isolated from the measuring apparatus.

[14] The ideas and prose of this Appendix are due entirely to Carlton M. Caves, and constitute a portion of the material submitted by him to the California Institute of Technology in partial fulfillment of the requirements for the Ph.D. degree.

Equations (D3) cannot be solved exactly with any ease, but when ε is small a good approximate solution can be obtained. The key to the approximation is the realization that the operators of Eqs. (D3) are nearly periodic with period π/ω. We implement the approximation by writing $\hat{X}_1$, $\hat{X}_2$, and $\hat{Q}$ as "Fourier series" with slowly varying "Fourier coefficients":

$$\hat{X}_1(t)=\sum_{n=-\infty}^{\infty}\hat{b}_n(t)e^{2in\omega t}, \tag{D4a}$$

$$\hat{X}_2(t)=\sum_{n=-\infty}^{\infty}\hat{c}_n(t)e^{2in\omega t}, \tag{D4b}$$

$$\hat{Q}(t)=\sum_{n=-\infty}^{\infty}\hat{d}_n(t)e^{2in\omega t}. \tag{D4c}$$

Of course, these expansions are "quasiunique" only for times greater than an oscillator period ($\omega t\gtrsim 2\pi$), but we are interested in the solutions only for such times. Hermiticity of $\hat{X}_1$, $\hat{X}_2$, and $\hat{Q}$ implies that $\hat{b}_{-n}=\hat{b}_n^\dagger$, etc., and as we show below, the assumption of slowly varying Fourier coefficients is satisfied because $d\hat{b}_n/dt\sim\varepsilon\omega\hat{b}_n$, etc. [see Eqs. (D7)].

To proceed, we plug the expansions (D4) into Eqs. (D3) and equate terms with the same rapid time dependence. The result is a set of coupled differential equations for the Fourier coefficients. We then simplify these equations by neglecting time derivatives in all equations except the $n=0$ equations—a step justified by the slowly varying character of the Fourier coefficients. The resulting coupled equations are all algebraic, except the $n=0$ equations.

Little would be gained by expansions (D4) if we had to consider all terms in the expansions. Fortunately, we need not do so. We are interested only in the largest terms in each expansion; and beyond the first term or two, each expansion becomes a power series in the small quantity ε. Indeed, using the coupled equations for the Fourier coefficients, one can easily show that, for $n\geq 1$ and for $\omega\tau\gg 1$, $\hat{b}_n\sim\varepsilon^{n-1}\hat{b}_0$, $\hat{c}_n\sim\varepsilon^n\hat{c}_0$, and $\hat{d}_n\sim\varepsilon^n\hat{d}_0$. Consequently, the only coefficients of interest are those with $n=0$ and $n=1$; and the $n=1$ terms can be neglected in the expansions for $\hat{X}_2$ and $\hat{Q}$, but they must be retained in the expansions for $\hat{X}_1$ and $\hat{\Pi}$. The $n=1$ equations can then be used to write the remaining $n=1$ coefficients in terms of $n=0$ coefficients. Putting all this together, one finds, at this level of approximation,

$$\hat{X}_1=\hat{b}_0-(K/4m\omega^2)\hat{d}_0\cos 2\omega t, \tag{D5a}$$

$$\hat{X}_2=\hat{c}_0, \tag{D5b}$$

$$\hat{Q}=\hat{d}_0, \tag{D5c}$$

$$\hat{\Pi}/L=(d\hat{d}_0/dt)+(K/4L\omega)\hat{c}_0\cos 2\omega t, \tag{D5d}$$

where the operators $\hat{b}_0$, $\hat{c}_0$, and $\hat{d}_0$ satisfy the coupled equations

$$d\hat{b}_0/dt=\tfrac{1}{4}\varepsilon^2\omega\hat{c}_0, \tag{D6a}$$

$$d\hat{c}_0/dt=-(K/2m\omega)\hat{d}_0, \tag{D6b}$$

$$d^2\hat{d}_0/dt^2=\varepsilon^2\omega^2\hat{d}_0-(K/2L)\hat{b}_0. \tag{D6c}$$

In Eqs. (D5d), (D6b), and (D6c) terms proportional to $\hat{b}_0$ and $\hat{d}_0$ have been omitted because they are negligible.

Equations (D6) can be solved easily. When the solutions are written in terms of appropriate initial values at $t=0$, they have the form

$$\begin{aligned}\sqrt{3}\hat{b}_0=&\hat{X}_1(0)(\nu_2^2\cosh\nu_1u+\nu_1^2\cos\nu_2u)+\tfrac{1}{4}\varepsilon\hat{X}_2(0)(\nu_1\sinh\nu_1u+\nu_2\sin\nu_2u)\\&+(K/8m\omega^2)\hat{Q}(0)[(2\nu_2^2-1)\cosh\nu_1u+(2\nu_1^2+1)\cos\nu_2u]\\&-(K/8mL\omega^3)\varepsilon^{-1}\hat{\Pi}(0)(\nu_1^{-1}\sinh\nu_1u-\nu_2^{-1}\sin\nu_2u),\end{aligned} \tag{D7a}$$

$$\begin{aligned}\sqrt{3}\hat{c}_0=&4\varepsilon^{-1}\hat{X}_1(0)(\nu_1\nu_2^2\sinh\nu_1u-\nu_2\nu_1^2\sin\nu_2u)+\hat{X}_2(0)(\nu_1^2\cosh\nu_1u+\nu_2^2\cos\nu_2u)\\&+(K/2m\omega^2)\varepsilon^{-1}\hat{Q}(0)[\nu_1(2\nu_2^2-1)\sinh\nu_1u-\nu_2(2\nu_1^2+1)\sin\nu_2u]\\&-(K/2mL\omega^3)\varepsilon^{-2}\hat{\Pi}(0)(\cosh\nu_1u-\cos\nu_2u),\end{aligned} \tag{D7b}$$

$$\begin{aligned}\sqrt{3}\hat{d}_0=&-(4m\omega^2/K)\hat{X}_1(0)(\cosh\nu_1u-\cos\nu_2u)-(2m\omega^2/K)\varepsilon\hat{X}_2(0)(\nu_1^3\sinh\nu_1u-\nu_2^3\sin\nu_2u)\\&+\hat{Q}(0)(\nu_2^2\cosh\nu_1u+\nu_1^2\cos\nu_2u)+\varepsilon^{-1}[\hat{\Pi}(0)/L\omega](\nu_1\sinh\nu_1u+\nu_2\sin\nu_2u),\end{aligned} \tag{D7c}$$

where $u\equiv\varepsilon\omega t$, and where

$$\nu_1\equiv(\sqrt{3}+1)^{1/2}/\sqrt{2},\quad \nu_2\equiv(\sqrt{3}-1)^{1/2}/\sqrt{2}. \tag{D8}$$

Note that the characteristic time scale of these solutions is $(\varepsilon\omega)^{-1}$, so the Fourier coefficients do indeed vary slowly in time. The reader might be bothered by the exponential instability of these solutions, but she should not be. As we show below, any real measurement will not last longer than a time $\tau\sim(\varepsilon\omega)^{-1}$.

Equations (D5) and (D7) give the free evolution of the coupled oscillator-meter system, and they can be applied to analyzing a measurement. The measurement process we consider is similar to that described in Sec. III.B and Appendix C. The measurement begins at $t=0$; the oscillator and meter interact via the interaction Hamiltonian (D1) for a time τ; and at the end of this time the readout system reads out a value for the meter coordinate, from which the experimenter infers a value for X_1.

To analyze the measurement, we must first specify the initial ($t=0$) states of the oscillator and meter. Our objective in this Appendix is to find the best possible performance of single-transducer back-action-evading measurements, so we shall choose the initial states to

optimize the measurement accuracy. We assume that at $t=0$ the oscillator is in a Gaussian (minimum-uncertainty) wave-packet state (in X_1) with $\langle\hat{X}_1(0)\rangle=\xi_0$ and $\langle\hat{X}_2(0)\rangle=0$. The meter is prepared in a Gaussian wave packet (in Q) with $\langle\hat{Q}(0)\rangle=\langle\hat{\Pi}(0)\rangle=0$. The initial variances $(\Delta X_1)_0$ and $(\Delta Q)_0$ are chosen to minimize the variance of the meter coordinate at $t=\tau$:

$$(\Delta X_1)_0^2=\varepsilon(\hbar/4m\omega)CD^{-1}, \tag{D9a}$$

$$(\Delta Q)_0^2=\varepsilon^{-1}(\hbar/2L\omega)AB^{-1}, \tag{D9b}$$

where

$$A\equiv\nu_1\sinh\varepsilon\omega\nu_1\tau+\nu_2\sin\varepsilon\omega\nu_2\tau, \tag{D10a}$$

$$B\equiv\nu_2^2\cosh\varepsilon\omega\nu_1\tau+\nu_1^2\cos\varepsilon\omega\nu_2\tau, \tag{D10b}$$

$$C\equiv\nu_1^3\sinh\varepsilon\omega\nu_1\tau-\nu_2^3\sin\varepsilon\omega\nu_2\tau, \tag{D10c}$$

$$D\equiv\cosh\varepsilon\omega\nu_1\tau-\cos\varepsilon\omega\nu_2\tau \tag{D10d}$$

[Eqs. (D5c) and (D7c)].

The oscillator and meter interact for a time τ, during which the expectation value of the meter coordinate gets displaced to

$$\langle\hat{Q}(\tau)\rangle=-(4m\omega^2/K\sqrt{3})\xi_0D, \tag{D11a}$$

and the variance of the meter coordinate grows to

$$\Delta Q(\tau)=\varepsilon^{-1/2}(\hbar/3L\omega)^{1/2}(AB+CD)^{1/2} \tag{D11b}$$

[Eqs. (D5c), (D7c), (D9), and (D10)]. At time τ the readout system reads out a value Q_m for the meter coordinate. Using Eq. (D11a) the experimenter infers a value

$$\xi_m=-(K\sqrt{3}/4m\omega^2)Q_mD^{-1} \tag{D12a}$$

for X_1. In a set of measurements on an ensemble of identical systems, the mean of this inferred value is ξ_0, and its variance is

$$\Delta\xi_m=\varepsilon^{1/2}(\hbar/2m\omega)^{1/2}(ABD^{-2}+CD^{-1})^{1/2}. \tag{D12b}$$

The measurement can determine ξ_0 with probable error $\Delta\xi_m$. In Eqs. (D11b) and (D12b) the first term on the right-hand side comes from the initial uncertainty in the meter coordinate and the second from the initial uncertainty in X_1.

3. Discussion

Interpretation of Eqs. (D11) and (D12) is obscured by their complicated dependence on τ. Their meaning is made a great deal clearer by looking at their form for short and long measurement times. For short measurement times ($\varepsilon\omega\tau\ll1$ but $\omega\tau\gg2\pi$) the meter displacement (D11a) and the probable error (D12b) are

$$\langle\hat{Q}(\tau)\rangle=-(K\tau^2/4L)\xi_0, \tag{D13a}$$

$$\Delta\xi_m=2\varepsilon^{-1}(\omega\tau)^{-3/2}(\hbar/2m\omega)^{1/2}=(16\hbar L/K^2\tau^3)^{1/2}. \tag{D13b}$$

[One can verify from Eqs. (D3) that these expressions are also valid to within factors of order unity when $\omega\tau\sim2\pi$.] The probable error (D13b) is due entirely to uncertainties in the meter; for short measurement times minimization of the uncertainty due to the initial oscillator variances is unimportant. Indeed, as long as $(\Delta X_1)_0^2$ is somewhat greater than its optimum value $[(\Delta X_1)_0>(\omega\tau)^{-1/2}(\hbar/2m\omega)^{1/2}$; cf. Eq. (D9a)], the probable error has the form

$$\Delta\xi_m\simeq[(16\hbar L/K^2\tau^3)+(\Delta X_1)_0^2]^{1/2}. \tag{D13c}$$

If this measurement of X_1 is to be repeatable to within the error $\Delta\xi_m$, then the condition $\Delta X_1(\tau)\lesssim\Delta\xi_m$ must be satisfied; otherwise, at the time of the readout of Q, the expectation value of $\hat{X}_1$ will "jump" an unknown distance greater than $\Delta\xi_m$. That this condition holds for $\tau\lesssim(\varepsilon\omega)^{-1}$ can be easily verified using Eqs. (D5a) and (D7).

The important feature of Eqs. (D13a) and (D13c) is that they are virtually identical to the comparable equations for measurements with perfect coupling to $\hat{X}_1$ [cf. Eqs. (3.19a) and (3.20b)]. The only difference is a factor of 2 in each equation; and this factor can be traced to the fact that, with the momentum coupling omitted, the mean force on the meter is cut in half [cf. Eqs. (3.17d) and (D3d)]. Just as in the case of perfect $\hat{X}_1$ coupling, single-transducer back-action-evading measurements beat the standard quantum limit when $\omega\tau\gtrsim\varepsilon^{-2/3}$. *Conclusion*: For short measurement times $\tau\ll(\varepsilon\omega)^{-1}$ (but $\tau\gtrsim\omega^{-1}$) the imperfection of coupling to $\hat{X}_1$ has no significant effect on the measurement accuracy, because there has not been time for "noise" in the measuring apparatus to "feed back" onto $\hat{X}_1$ and disturb it significantly.

The difference between single-transducer and perfectly coupled two-transducer back-action-evading measurements shows up at long measurement times ($\varepsilon\omega\tau\gg1$), when the meter displacement (D11a) and measurement error (D12b) become

$$\langle\hat{Q}(\tau)\rangle=-(2m\omega^2/K\sqrt{3})\xi_0e^{\varepsilon\omega\nu_1\tau}, \tag{D14a}$$

$$\Delta\xi_m=(\varepsilon\nu_1\sqrt{3})^{1/2}(\hbar/2m\omega)^{1/2}. \tag{D14b}$$

In the case of a perfectly coupled measurement, one can choose the initial variance of $\hat{X}_1$ as small as desired (in principle); then the measurement becomes more and more accurate as τ increases [Eq. (3.20b)]. However, for a single-transducer measurement, the accuracy does not continue to improve; instead it hits a "floor" at approximately $\varepsilon^{1/2}(\hbar/2m\omega)^{1/2}$ for times $\tau\gtrsim(\varepsilon\omega)^{-1}$, because $\hat{X}_1$ no longer successfully evades "back-action" noise from the measuring apparatus. Note that for long measurement times $\tau\gg(\varepsilon\omega)^{-1}$, the measurements are not repeatable because $\Delta X_1(\tau)\gg\Delta\xi_m$ [cf. Eqs. (D5a) and (D7)].

The dependence of the measurement accuracy (D12b) on τ can be conveniently summarized by using only the small and large τ forms:

$$\Delta\xi_m\simeq\begin{cases}\varepsilon^{-1}(\omega\tau)^{-3/2}(\hbar/2m\omega)^{1/2}, & \varepsilon\lesssim\varepsilon\omega\tau\lesssim1, \quad \text{(D15a)}\\ \varepsilon^{1/2}(\hbar/2m\omega)^{1/2}, & \varepsilon\omega\tau\gtrsim1. \quad \text{(D15b)}\end{cases}$$

This behavior is similar to what one expects for amplitude-and-phase measurements. The accuracy of an amplitude-and-phase measurement should improve as τ increases, but it must eventually hit a floor at the standard quantum limit $(\hbar/2m\omega)^{1/2}$. The floor must be at the standard quantum limit because X_1 and X_2 are measured with equal precision, so they are equally affected by back-action noise. The accuracy floor for single-transducer back-action-evading measurements is lower because X_1 is partially shielded from back-action noise.

Up to now we have operated under the assumption that the coupling constant ε is fixed, and we have investigated the dependence of the measurement error on τ for fixed ε. One can adopt a different point of view—that the coupling strength is under the control of the experimenter. Given this freedom, the experimenter will choose the value of ε (by choosing L) to optimize the measurement accuracy for a given measurement time τ ($\gtrsim \omega^{-1}$). The choice he will make is $\varepsilon \simeq (\omega\tau)^{-1}$, and the measurement error (D15) will be

$$\Delta\xi_m \simeq (\omega\tau)^{-1/2}(\hbar/2m\omega)^{1/2}. \tag{D16}$$

This is the optimum performance for a single-transducer measurement of the simple type considered in this Appendix [cf. Eq. (2.43)].

So far in this Appendix we have considered a meter with no "restoring force." In practice this is not usually the case; in a typical design such as that in Appendix B.1, the meter is an LC circuit (term $\hat{Q}^2/2C$ added to meter Hamiltonian, where C is the total capacitance in the circuit including that associated with the position transducer). In this situation the analysis given in this Appendix will apply approximately for measurement times smaller than the characteristic time of the circuit—i.e., $\tau \lesssim \tilde{\tau} \equiv (LC)^{1/2}$. If $\tilde{\tau} \gtrsim (\varepsilon\omega)^{-1}$, i.e., $K^2C/8m\omega^2 \gtrsim 1$, the effect of the capacitor can essentially be ignored, because the preceding analysis applies for times long enough to hit the accuracy floor. However, in practice it may be difficult to make the capacitance large enough, and one may be stuck with the case $\tilde{\tau} \ll (\varepsilon\omega)^{-1}$, i.e., $K^2C/8m\omega^2 \ll 1$.

To analyze this case in detail requires a more sophisticated model for the measuring apparatus than we have used here. We consider more sophisticated measuring systems in Paper II, and we analyze their performance using semiclassical techniques. However, we can get a good idea of the potential performance from the preceding analysis.

A measurement of duration $\tilde{\tau}$ has accuracy $\simeq \varepsilon^{-1}(\omega\tilde{\tau})^{-3/2}(\hbar/2m\omega)^{1/2}$. A measurement of duration $\tau \gtrsim \tilde{\tau}$ can be regarded as a sequence of measurements of duration $\tilde{\tau}$. Before the initial measurement in the sequence, the oscillator is prepared in a Gaussian wave-packet state (in X_1). Appendix C analyzed a sequence of measurements with perfect coupling to $\hat{X}_1$. In that analysis the variance of $\hat{X}_1$ always decreased during the sequence. Here, with imperfect coupling, we expect the variance of $\hat{X}_1$ to decrease until it is approximately equal to the optimum value for measurements of duration $\tilde{\tau}$: $(\Delta X_1)_0 \simeq (\omega\tilde{\tau})^{-1/2}(\hbar/2m\omega)^{1/2}$ [cf. Eq. (D9a)]. Thus we shall choose the initial variance of $\hat{X}_1$ to be this optimum value; then the variance should not change significantly during the sequence.

The results of all the measurements in the sequence are used to determine the *initial* expectation value of $\hat{X}_1$. The accuracy of this determination improves as the square root of the number of measurements. Thus the measurement error for a sequence of total duration $\tau \gtrsim \tilde{\tau}$ is given approximately by

$$\Delta\xi \simeq \varepsilon^{-1}(\omega\tilde{\tau})^{-3/2}(\hbar/2m\omega)^{1/2}(\tilde{\tau}/\tau)^{1/2} \simeq (\beta\omega\tau)^{-1/2}(\hbar/2m\omega)^{1/2}, \tag{D17}$$

where β is a dimensionless coupling constant defined by

$$\beta \equiv (\varepsilon\omega\tilde{\tau})^2 = K^2C/8m\omega^2. \tag{D18}$$

The improvement in accuracy of Eq. (D17) does not continue forever, because the expectation value of $\hat{X}_1$ changes during the sequence. In particular, the expectation value of $\hat{X}_1$ "jumps" at the time of each readout; the detailed analysis in Appendix C suggests that the expected magnitude of each jump is approximately

$$\frac{(\Delta X_1)_0^2}{\varepsilon^{-1}(\omega\tilde{\tau})^{-3/2}(\hbar/2m\omega)^{1/2}} \simeq \varepsilon(\omega\tilde{\tau})^{1/2}\left(\frac{\hbar}{2m\omega}\right)^{1/2}$$

[see Eq. (C36) and accompanying discussion]. {The expectation value of $\hat{X}_1$ also changes *during* each measurement in the sequence because of the imperfect coupling to $\hat{X}_1$ [Eqs. (D5a) and (D7)], but these changes are negligible for $\tilde{\tau} \ll (\varepsilon\omega)^{-1}$, provided that one uses a "feedback force" on the meter like that in Appendix C.} The jumps add randomly, so that after a time $\tau \gtrsim \tilde{\tau}$, the expectation value of $\hat{X}_1$ will have wandered a distance $\simeq \varepsilon(\omega\tau)^{1/2}(\hbar/2m\omega)^{1/2}$. The measurement accuracy improves as in Eq. (D17) only until the distance wandered becomes comparable to the measurement error. Thus the accuracy hits a floor at approximately $(\omega\tilde{\tau})^{-1/2}(\hbar/2m\omega)^{1/2}$ for measurement times $\tau \gtrsim \tilde{\tau}/\beta$. The accuracy floor is approximately equal to the initial variance of $\hat{X}_1$—i.e., the entire sequence allows one to determine the initial expectation value of $\hat{X}_1$ with an error of order the initial variance.

The dependence of measurement error on τ can be summarized as follows:

$$\Delta\xi \simeq \begin{cases} (\beta\omega\tau)^{-1/2}(\tilde{\tau}/\tau)(\hbar/2m\omega)^{1/2}, & \tau \lesssim \tilde{\tau}, \quad \text{(D19a)} \\ (\beta\omega\tau)^{-1/2}(\hbar/2m\omega)^{1/2}, & \tilde{\tau} \lesssim \tau \lesssim \tilde{\tau}/\beta, \quad \text{(D19b)} \\ (\omega\tilde{\tau})^{-1/2}(\hbar/2m\omega)^{1/2}, & \tau \gtrsim \tilde{\tau}/\beta. \quad \text{(D19c)} \end{cases}$$

Note that Eqs. (D19) simplify to Eqs. (D15) when $\beta \simeq 1$.

Just as in the previous case ($\beta \gtrsim 1$), so in this case ($\beta \lesssim 1$), the optimum performance is achieved by adjusting $\tilde{\tau}$ (adjusting L) to obtain the best accuracy for a given τ. The optimum choice is $\beta\tau \lesssim \tilde{\tau} \lesssim \tau$, and the resulting optimum accuracy is

$$\Delta\xi \simeq (\beta\omega\tau)^{-1/2}(\hbar/2m\omega)^{1/2}. \tag{D20}$$

It should now be clear that β is the really important measure of coupling strength for this type of single-transducer back-action-evading measurement. For $\beta \gtrsim 1$ the optimum performance is given by Eq. (D16); for $\beta \lesssim 1$, by Eq. (D20).

The constant β is (to within factors of order unity) a Gibbons-Hawking (1971) coupling constant. In Paper II we give an exact definition of the Gibbons-Hawking constant for an arbitrary measuring system coupled to an oscillator; we present a semiclassical derivation of the limiting accuracy (D20) for such a system; and we generalize that accuracy to the case where the system contains an amplifier with noise temperature greater than the quantum limit ($\hbar \to 2kT_n/\omega$); cf. Braginsky, Vorontsov, and Khalili (1978), Thorne *et al.* (1979), and Braginsky *et al.* (1980).

In deriving the optimum performances for strong ($\beta \gtrsim 1$) and weak ($\beta \lesssim 1$) coupling [Eqs. (D16) and (D20)], we assumed that $\tilde{\tau}$ (or equivalently L) can be adjusted so as to match the measurement time τ. It is important to remember that, if $\tilde{\tau}$ is fixed by practical con-

siderations, then a measurement with either strong or weak coupling will hit an accuracy floor

$$\Delta\xi \simeq (\omega\tilde{\tau})^{-1/2}(\hbar/2m\omega)^{1/2} \tag{D21}$$

as τ increases [cf. Eq. (D19c)]. This accuracy floor is an absolute limit for continuous single-transducer measurements. For a realistic continuous single-transducer measurement, $\tilde{\tau}$ is the averaging time of the filter which precedes the amplifier and which averages the modulated transducer output (see Sec. II.F.3 and Paper II).

REFERENCES

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

Aharonov, Y., and D. Bohm, 1964, Phys. Rev. **134**, 1417B.

Aharonov, Y., and A. Petersen, 1971, in *Quantum Theory and Beyond*, edited by T. Bastin (Cambridge University, Cambridge, England), p. 135.

Allen, M. A., Z. D. Farkas, H. A. Hogg, E. W. Hoyt, and P. B. Wilson, 1971, IEEE Trans. Nucl. Sci. **18**, 168.

Bagdasarov, Kh. S., V. B. Braginsky, V. P. Mitrofonov, and V. S. Shiyan, 1977, Vestnik Moskovskovo Universiteta, Ser. Fiz. Astr. **1**, 98.

Bertotti, B., 1977, editor, *Gravitazione Sperimentale*, Proceedings of an international conference held at Pavia, Italy, 17–20 September 1976 (Accademia Nazionale dei Lincei, Rome).

Bohm, D., 1951, *Quantum Theory* (Prentice-Hall, Englewood Cliffs), especially Chaps. 6 and 22.

Bohr, N., 1928, Naturwissenschaften **16**, 245.

Braginsky, V. B., 1970, *Physical Experiments with Test Bodies* (Nauka, Moscow) [English translation published as NASA-TT F672, National Technical Information Service, Springfield, VA]; especially Eqs. (3.17) and (3.25).

Braginsky, V. B., 1974, in *Gravitational Radiation and Gravitational Collapse*, Proceedings of IAU Symposium No. 64, edited by Cécile DeWitt-Morette (Reidel, Dordrecht), p. 28.

Braginsky, V. B., 1977, in *Topics in Theoretical and Experimental Gravitation Physics*, edited by V. De Sabbata and J. Weber (Plenum, New York), p. 105.

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

Braginsky, V. B., L. P. Grishchuk, A. G. Doroshkevich, Ya. B. Zel'dovich, I. D. Novikov, and M. V. Sazhin, 1973, Zh. Eksp. Teor. Fiz. **65**, 1729 [Sov. Phys. JETP **38**, 865 (1974)].

Braginsky, V. B., V. I. Panov, V. G. Petnikov, and V. D. Popel'nyuk, 1977, Prib. Tekh. Eksp. **20**(1), 234 [Instrum. Exp. Tech. (USSR) **20**, 269 (1977)].

Braginsky, V. B., and V. N. Rudenko, 1978, Phys. Rep. **46**, 165.

Braginsky, V. B., and Yu. I. Vorontsov, 1974, Usp. Fiz. Nauk **114**, 41 [Sov. Phys.—Usp. **17**, 644 (1975)].

Braginsky, V. B., Yu. I. Vorontsov, and F. Ya. Khalili, 1977, Zh. Eksp. Teor. Fiz. **73**, 1340 [Sov. Phys. JETP **46**, 705 (1977)].

Braginsky, V. B., Yu. I. Vorontsov, and F. Ya. Khalili, 1978, Zh. Eksp. Teor. Fiz. Pis'ma Red. **27**, 296 [JETP Lett. **27**, 276 (1978)].

Braginsky, V. B., Yu. I. Vorontsov, and V. D. Krivchenkov, 1975, Zh. Eksp. Teor. Fiz. **68**, 55 [Sov. Phys. JETP **41**, 28 (1975)].

Braginsky, V. B., Yu. I. Vorontsov, and K. S. Thorne, 1980, Science, in press.

Carruthers, P., and M. M. Nieto, 1965, Phys. Rev. Lett. **14**, 387.

Caves, C. M., 1979, Phys. Lett. **B80**, 323.

Douglass, D. H., and V. B. Braginsky, 1979, in *General Relativity: An Einstein Centenary Survey*, edited by S. W. Hawking and W. Israel (Cambridge University, Cambridge, England), p. 90.

Drever, R. W. P., J. Hough, W. A. Edelstein, J. R. Pugh, and W. Martin, 1977, pp. 365–369 of Bertotti (1977).

Edelstein, W. A., J. Hough, J. R. Pugh, and W. Martin, 1978, J. Phys. E **11**, 710.

Epstein, R., and J. P. A. Clark, 1979, in *Sources of Gravitational Radiation*, edited by L. Smarr (Cambridge University, Cambridge, England), p. 477.

Fock, V., 1962, Zh. Eksp. Teor. Fiz. 42, 1135 [Sov. Phys. JETP **15**, 784 (1962)].

Gibbons, G. W., and S. W. Hawking, 1971, Phys. Rev. D **4**, 2191.

Giffard, R., 1976, Phys. Rev. D **14**, 2478.

Grishchuk, L. P., and M. V. Sazhin, 1975, Zh. Eksp. Teor. Fiz. **68**, 1569 [Sov. Phys. JETP **41**, 787 (1975)].

Heffner, H., 1962, Proc. IRE **50**, 1604.

Hollenhorst, J. N., 1979, Phys. Rev. D **19**, 1669.

McGuigan, D. F., C. C. Lam, R. Q. Gram, A. W. Hoffmann, D. H. Douglass, and H. W. Gutche, 1978, J. Low Temp. Phys. **30**, 621.

Merzbacher, E., 1970, *Quantum Mechanics*, 2nd ed. (Wiley, New York), especially Secs. 15.8–15.10.

Misner, C. W., K. S. Thorne, and J. A. Wheeler, 1973, *Gravitation* (Freeman, San Franciso).

Pauli, W., 1958, in *Handbuch der Physik*, Vol. 5, Part 1, edited by S. Flügge (Springer, Berlin), especially pp. 73–74.

Pegoraro, F., E. Picasso, and L. A. Radicati, 1978, J. Phys. A **11**, 1949.

Pfister, H., 1976, Cryogenics **16**, 17.

Robinson, F. N. H., 1974, *Noise and Fluctuations in Electronic Devices and Circuits* (Clarendon, Oxford).

Serber, R., and C. H. Townes, 1960, in *Quantum Electronics, a Symposium*, edited by C. H. Townes (Columbia University, New York), p. 233.

Smagin, A. G., 1974, Prib. Tekh. Eksp. **17**(6), 143 [Instrum. Exp. Tech. (USSR) **17**, 1721 (1974)].

Takahasi, H., 1965, Adv. Commun. Syst. **1**, 227, especially Sec. XI.

Thorne, K. S., 1978, in *Theoretical Principles in Astrophysics and Relativity*, edited by N. R. Lebovitz, W. H. Reid, and P. O. Vandervoort (University of Chicago, Chicago), especially pp. 163–174.

Thorne, K. S., 1980, Rev. Mod. Phys. **52**, 285, 299.

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

Thorne, K. S., C. M. Caves, V. D. Sandberg, M. Zimmermann, and R. W. P. Drever, 1979, in *Sources of Gravitational Radiation*, edited by L. Smarr (Cambridge University, Cambridge, England), p. 49.

Tyson, J. A., and R. P. Giffard, 1978, Ann. Rev. Astron. Astrophys. **16**, 521.

Unruh, W. G., 1977, preprint from University of British Columbia; published in revised form as Unruh (1978).

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

Unruh, W. G., 1979, Phys. Rev. D **19**, 2888.

Unruh, W. G., 1980, to be published in *Gravitational Radiation and Collapsed Objects*, edited by C. Edwards (Springer, Berlin).

Von Neumann, J., 1932, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin) {English translation: *Mathematical Foundations of Quantum Mechanics* [Princeton University, Princeton, NJ (1955)]}.

Wagoner, R. V., C. M. Will, and H.-J. Paik, 1979, Phys. Rev. D **19**, 2325.

Weber, J., 1959, Rev. Mod. Phys. **31**, 681; see especially Sec. V.

Weber, J., 1979, in *Einstein Centenary Volume*, edited by A. Held (Plenum, New York).

Weiss, R., 1972, Progress Report 105, Res. Lab. Electron., MIT, p. 54.

Weiss, R., 1979, in *Sources of Gravitational Radiation*, edited by L. Smarr (Cambridge University, Cambridge, England), p. 7.

Wyler, J. A., 1974, J. Gen. Rel. Grav. **5**, 175.