AAPT PHYSICS EDUCATION
AMERICAN JOURNAL of PHYSICS

Uncertainty principle for joint measurement of noncommuting variables

M. G. Raymer

Citation: American Journal of Physics **62**, 986 (1994); doi: 10.1119/1.17657
View online: http://dx.doi.org/10.1119/1.17657
View Table of Contents: http://scitation.aip.org/content/aapt/journal/ajp/62/11?ver=pdfcov
Published by the American Association of Physics Teachers

Uncertainty principle for joint measurement of noncommuting variables

M. G. Raymer[a)]

Department of Physics and Chemical Physics Institute, University of Oregon, Eugene, Oregon 97403

(Received 27 December 1993; accepted 12 May 1994)

Does the Heisenberg uncertainty principle refer to errors of measurement, or to the spread of values of the physical variables intrinsic to a particle's state, or to some combination of these? The most commonly quoted form of the uncertainty principle relates the spread of an ensemble of separate measurements of some variable to the analogous spread of its conjugate variable. In contrast, Heisenberg's original argument for the uncertainty principle involved the perturbation to a particle's state by a measurement of one variable, which affects one's ability to predict the outcome of a subsequent measurement of the conjugate variable. The relation between these two views of the uncertainty principle is discussed in this paper. A familiar example is considered: an ensemble of identically prepared particles passing through a slit, and after further propagation being detected. From this arrangement it is possible to infer joint (although necessarily imprecise) information on *both* transverse position and conjugate momentum for each member of the ensemble. It is shown that in this case of joint measurement the product of standard deviations for the measurement outcomes is at least twice as large as the lower bound implied by the usual uncertainty principle. The discussion is meant to help clarify the different roles played in the various statements of the uncertainty principle by the initial state and by measurement error.

I. INTRODUCTION

Recently, with the advent of ultraprecise measurement techniques being developed in quantum optics,[1] atomic physics,[2] and in gravitational wave detection,[3] among other areas, renewed interest has arisen concerning the ultimate limitations of measurement imposed by quantum mechanics. An interesting question concerns the limits placed on the *simultaneous* (or *joint*) measurability of noncommuting variables. In optics these might be the real and imaginary parts of the field (i.e., in-phase and in-quadrature amplitudes), or they might be the optical phase and the photon number. In atomic physics these might be the position and momentum of an atom in a beam or a trap. In gravitational wave detection these might be the position and momentum of a mechanically resonant bar antenna. An important consequence of quantum mechanics, which does not seem to be widely known, is that such joint measurements invariably introduce an uncertainty that is beyond that associated with the usual Heisenberg uncertainty principle. This was apparently first pointed out in the modern context by Arthurs and Kelly in 1965,[4] and has been recognized as fundamentally important by current researchers, as reviewed by Yamamoto and Haus,[5] and by Stenholm.[6]

The "usual" (most commonly quoted) uncertainty principle states that the uncertainty product obtained for separate measurements of position *or* momentum is $\Delta x \, \Delta p_x \geqslant \hbar/2$, where the uncertainties Δx and Δp_x are defined as standard deviations of an ensemble of measurements made on a system in a given state. The usual principle refers to the intrinsic spread of the variables in the quantum state of the system, due to their quantum indeterminacy, and is not a consequence of any measurement error.

The extended uncertainty principle of Arthurs and Kelly states that if joint (although necessarily imprecise) measurements are made on *both* position x and conjugate momentum p_x for each member of an ensemble of identically prepared systems, the product of standard deviations for the measurement outcomes x_M and p_{xM} satisfies $\Delta x_M \Delta p_{xM} \geqslant \hbar$. It is at least twice as large as the lower bound implied by the usual uncertainty principle. The extended principle accounts for the added measurement error, or noise, that is a result of attempting to measure (or at least estimate) the values of two noncommuting variables for each ensemble member. This extended principle is entirely a consequence of standard quantum mechanics, and one purpose of this paper is to demonstrate this result in the context of a standard example found in many elementary textbooks—a particle traveling through a slit.

The present discussion is meant to help clarify a common confusion that arises in many textbooks, and which can be traced to the original papers of Heisenberg:[7,8] namely whether the uncertainty principle refers to errors of measurement or to the spread of values of the physical variables intrinsic to the state, or to some combination of these. The history of this confusion has been reviewed and interpreted by Hilgevoord and Uffink.[9] It is demonstrated here that there are actually two different uncertainty principles, each valid in different cases, distinguished by the manner in which the data are collected—whether by separate or by joint measurements. This is consistent with the discussion of Arthurs and Kelly.

The original argument of Arthurs and Kelly[4] is actually very simple mathematically, but is somewhat abstract physically, and the present example is meant to be more concrete. The essence of their argument is this—the two noncommuting variables must each be made to interact with a different "meter" variable, and these two meter variables must themselves commute in order to be jointly measurable with arbitrarily high precision. It is in the coupling of the two original variables to the meter variables that the additional noise is introduced.

First a concrete example will be given and analyzed qualitatively: an ensemble of identically prepared particles passing through a slit, and after further propagation being detected. From this arrangement it is possible to infer joint (although necessarily imprecise) information on *both* position and conjugate momentum for each member of the ensemble. The limitations of this model will be discussed as a motivation for a more powerful treatment of a more general class of measurements, using recently developed generalizations of von Neumann's theory of measurement. A somewhat more general statement of the joint-measurement uncertainty

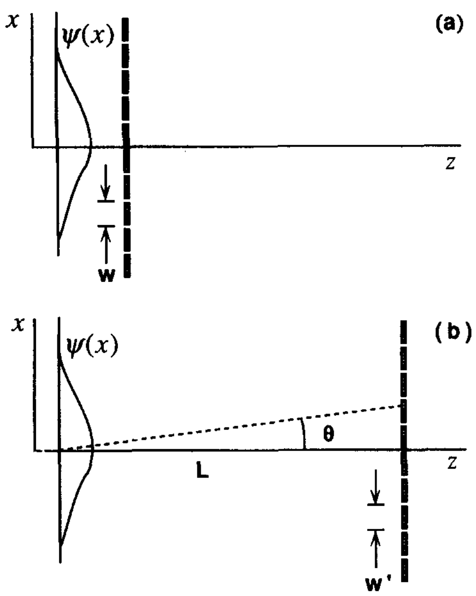

Fig. 1. Setup for *separate measurement* of the distribution of (a) transverse position or (b) transverse momentum for an ensemble of particles prepared in state $\psi(x)$ and nominally traveling in direction z. Detector plates are assumed to register the arrival of a single particle.

principle will be derived. The connection of this treatment to other discussions, notably those of Arthurs and Kelly,[4] Braginsky and Khalili,[10] and Lai and Haus[11] will be pointed out.

II. AN EXAMPLE: PARTICLE PASSING THROUGH A SLIT

A. Precise measurement—intrinsic indeterminacy relation

As an example consider an ensemble of identically prepared particles (say electrons) traveling nominally in the z direction, and partially localized in the transverse (x) direction with transverse distribution determined by the wave function $\psi(x)$, which is the same for all particles. It is desired to obtain information about the distribution of both position and momentum in this state, which is unknown to the experimenter. As shown in Fig. 1, the distribution $|\psi(x)|^2$ of particle positions in this state can be precisely determined by using an array of N detectors, each with width w, and counting the fraction of trials in which a particle is detected by each detector. Each detector may be represented as a metal plate connected to ground through a current meter which can register the passage of a single electron.

First assume that each detector is of negligible width, and the array fills the region accessible to the particles. Then the mean $\bar{x}_\psi$ and variance Δx_ψ^2 of the position measurements are given by the usual expressions[12]

$$\bar{x}_\psi = \int_{-\infty}^{\infty} x|\psi(x)|^2 dx, \tag{1a}$$

$$\Delta x_\psi^2 = \int_{-\infty}^{\infty} (x-\bar{x}_\psi)^2 |\psi(x)|^2 dx, \tag{1b}$$

which are intrinsic to the state ψ. The spread of position measurements here has nothing to do with errors in measurement, since the measurements can be made arbitrarily precise by making w smaller. Rather it reflects the physical indeterminacy of the particle's position.

Alternatively, the transverse momentum distribution intrinsic to the state can be measured by moving the detector array a sufficiently large distance L in the z direction and performing the same measurement (again with negligible detector width w'). For simplicity of notation I will denote the transverse momentum (conjugate to x) by p. We may assume that we know beforehand the particle's total momentum p_T (where $p_T \gg p$). Then if the particle is registered by the detector whose location makes an angle θ with the z axis [see Fig. 1(b)], its transverse momentum is known to be $p = p_T \sin\theta$. The transverse momentum wave function $\tilde{\psi}(p)$ is given by the Fourier transform of the position wave function,

$$\tilde{\psi}(p) = (2\pi\hbar)^{-1/2} \int_{-\infty}^{\infty} \psi(x) \exp(-ipx/\hbar) dx \tag{2}$$

and the transverse momentum distribution is given by $|\tilde{\psi}(p)|^2$. (This mapping of the momentum distribution into position for large L is analogous to far-field diffraction in optics.) The intrinsic mean $\bar{p}_\psi$ and variance Δp_ψ^2 of the transverse momentum are defined similarly to those for position in Eqs. (1), and these can be measured by the described method.

The usual uncertainty principle follows from these definitions and the properties of the Fourier transform (or equivalently the operator commutator $[\hat{x},\hat{p}] = i\hbar$), and states that the product of the measured standard deviations (square root of variance) satisfies[12]

$$\Delta x_\psi \Delta p_\psi \geq \hbar/2, \tag{3}$$

as was first proved by Kennard.[13] Again, these uncertainties are intrinsic to the incident state and have nothing to do with measurement error, as can be seen by the example just presented. I emphasize that in order to measure the quantities appearing in this relation, it is necessary to measure precisely one variable (position or momentum) only for each particle in the ensemble. Thus Eq. (3) refers to separate, rather than joint, measurement.

B. Imprecise measurement

If the detectors used in Fig. 1(a) have non-negligible widths w, and the jth detector is centered at position $X_j = jw (j = 1,2,\ldots,N)$, then the measured distribution of position will be broader than before, and can be calculated as follows. The probability for the particle to be registered by the jth detector is

$$P_D(X_j) = \int_{-\infty}^{\infty} h(X_j - x)|\psi(x)|^2 dx, \tag{4}$$

where the top-hat function is

$$h(X_j - x) = \begin{cases} 1, & -w/2 < (x - X_j) < w/2 \\ 0, & \text{otherwise} \end{cases}. \tag{5}$$

Because the particle must hit one of the detectors, these probabilities are normalized according to

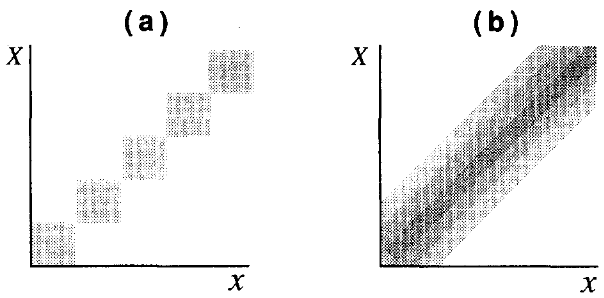

Fig. 2. Conditional probability, or resolution function, $f^2(X|x)$ for (a) sharp-edged, nonoverlapping detectors [Eq. (9)], or (b) smooth-edged, overlapping detectors [Eq. (20)]. This function is the probability that, given the particle is well localized at position x prior to measurement, the (imprecise) measurement will yield a result X.

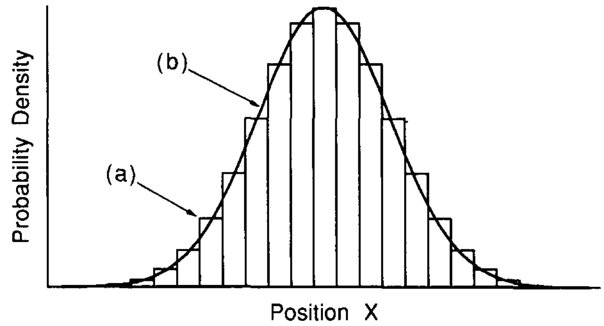

Fig. 3. (a) Measured distribution $P_{\psi f}(X)$ of position [Eq. (8)] and (b) corresponding smooth distribution $|\psi(X)|^2$ for the state $\psi(X)$.

$$\sum_{j=1}^{N} P_D(X_j)=1. \tag{6}$$

If the particle is registered by the jth detector, the experimenter can infer only that it was equally likely to have been between $(j-1/2)w$ and $(j+1/2)w$ prior to detection. This backward inference from an observation (Bayesian inference)[14] can be expressed as the conditional probability

$$P(X|X_j)=w^{-1}h(X_j-X) \tag{7}$$

that, given the particle was registered by the detector centered at X_j, its actual position (in a classical way of thinking) was X. The probability that the particle will be inferred from a measurement to have been at X is thus the sum of contributions from all detectors,

$$P_{\psi f}(X)=\sum_{j=1}^{N} P(X|X_j)P_D(X_j) \tag{8a}$$

$$=\int_{-\infty}^{\infty} f^2(X|x)|\psi(x)|^2dx, \tag{8b}$$

where $f^2(X|x)$ is the conditional probability that given the particle was actually at x prior to measurement, it will be inferred to have been at X. It follows from Eqs. (8) and (7) that its form is

$$f^2(X|x)=w^{-1}\sum_{j=1}^{N} h(X_j-X)h(X_j-x). \tag{9}$$

This "resolution function" is represented in Fig. 2(a), where it is seen to be equal to w^{-1} inside the squares along the diagonal and zero elsewhere. In words, this function says that if the particle is initially at a position x within the jth detector, it will be inferred equally likely to have been *anywhere* within this detector.

Equation (8) is the distribution of X that will be measured (inferred) according to this prescription if the particle is in state $\psi(x)$ and the measuring apparatus is described by f. It is a formal way to express that the smooth distribution $|\psi(x)|^2$ has been converted into a histogram, illustrated in Fig. 3. Using this histogram-type distribution, the mean value of X is given by

$$\bar{X}_{\psi f}=\int_{-\infty}^{\infty} XP_{\psi f}(X)dX=\sum_{j=1}^{N} X_jP_D(X_j), \tag{10}$$

where use was made of Eqs. (8) and (7). The variance of X is given by

$$\Delta X_{\psi f}^2=\int_{-\infty}^{\infty}(X-\bar{X}_{\psi f})^2P_{\psi f}(X)dX. \tag{11}$$

The standard deviation, the square root of the variance, is calculated using Eqs. (5)–(8), yielding

$$\Delta X_{\psi f}=\left(\sum_{j=1}^{N} P_D(X_j)(X_j-\bar{X}_{\psi f})^2+\Delta X_f^2\right)^{1/2} \tag{12a}$$

$$\cong\sqrt{\Delta x_\psi^2+\Delta X_f^2}, \tag{12b}$$

where

$$\Delta X_f^2=w^2/12 \tag{13}$$

is the X variance of the function $f^2(X|x)$ with x held constant, which makes it a simple top-hat function. The first term in Eq. (12a) is a discrete version of the intrinsic state variance Δx_ψ^2 in Eq. (1), and equals Δx_ψ^2 in the limit $w\rightarrow 0$. The form for the standard deviation in Eq. (12) can be understood as follows. First consider the limit in which the width of the wave function in x is much narrower than the detector widths w. Then, according to Eqs. (8) and (9), the distribution of X is a single top-hat function with width w and variance $\Delta X_f^2=w^2/12$. The opposite limit is simple—it corresponds to arbitrarily narrow detector widths, so the variance Δx_ψ^2 of the wave function dominates. The approximate form of Eq. (12b) is chosen as an estimate which interpolates smoothly between these two limits.

A similar argument may be developed for the measurement errors for the momentum, but this will not be used.

C. State reduction, or backaction influence of imprecise measurement

In order to gain some information about *both* position and momentum of a *single* particle in the incident state $\psi(x)$, we can modify the apparatus in Fig. 1 as follows. Remove one of the detector plates, say the one at X_J. Then given that we know the particle has impinged on the detector array, if it is not detected by one of the remaining detectors, we know that

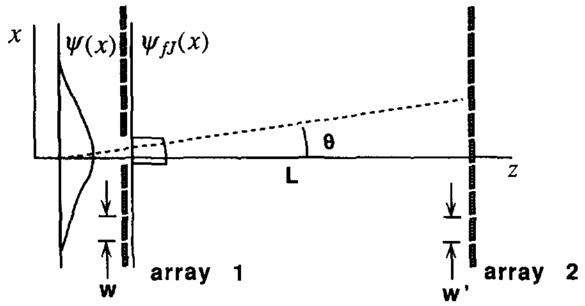

Fig. 4. Setup for *joint, imprecise measurement* of transverse position and momentum for each particle in an ensemble, all prepared in state $\psi(x)$. One detector plate in array 1 is removed, allowing the particle to pass through and travel to array 2 where it is registered.

it must have passed through the opening (or slit), with width w. This gives position information. As shown in Fig. 4, momentum information is then obtained by use of a second detector array located a large distance from the first.

Given that the particle has passed through the slit at X_J, the transmitted wave function is a truncated version of the incident wave function $\psi(x)$. According to the standard quantum measurement model of von Neumann (based on state projection, or reduction), the state becomes[15–17]

$$\psi_{fJ}(x)=\frac{h(X_J-x)\psi(x)}{\sqrt{\int h^2(X_J-x)|\psi(x)|^2dx}}, \tag{14}$$

where the denominator serves to renormalize the state. (Again we may make an optical analogy—the wave function behaves like the electric field of a light beam passing through an opaque aperture.) Notice that Eq. (14) is a projection—if we perform it again on $\psi_{fJ}(x)$ the state is reproduced.

D. Joint, imprecise measurement

Recall that our goal is to infer information about the incident state $\psi(x)$ by making joint, but imprecise, measurements of position and momentum on each individual particle taken from an ensemble of similarly prepared particles. In Fig. 4, passage through the slit in array 1 provides position information and detection at array 2 provides momentum information. We do not want the opening slit in array 1 to be too broad or no information about position will be obtained. We also do not want it to be too narrow or the transmitted state $\psi_{fJ}(x)$ in Eq. (14) will contain virtually no information about the transverse momentum of the incident state. An optimum slit width should be found.

Detector array 2, with negligible detector widths w', will yield precise measurements of transverse momentum on the transmitted state $\psi_{fJ}(x)$. The distribution of measured momenta will be given by $|\tilde{\psi}_{fJ}(p)|^2$ where $\tilde{\psi}_{fJ}(p)$ is the Fourier transform of $\psi_{fJ}(x)$. Unfortunately this distribution has an infinite variance, because the Fourier transform $\tilde{h}(p)$ of the top-hat function $h(x)$ behaves as $1/p$ for large p. (See the Appendix.) This makes it impossible in this example to express an uncertainty relation in terms of standard deviations.

Here we will overcome this problem in a simple qualitative way, with a more rigorous approach being discussed in Sec. III. Here we will use the full width $\delta p_{\psi f}$ of the distribution $|\tilde{\psi}_{fJ}(p)|^2$ instead of its variance as a measure of the spread in p. (It has actually been argued that full width is a better measure in general.)[9] Very roughly, this full width is given by (see the Appendix)

$$\delta p_{\psi f}\cong\sqrt{\Delta p_\psi^2+\delta p_f^2}, \tag{15}$$

where the first term is the intrinsic momentum spread in the incident state, and the second term, where

$$\delta p_f=2\pi\hbar/w, \tag{16}$$

is from the spreading of the wave function caused by the effect of the slit. (Again this result can be understood with a optical analogy—the slit causes diffraction.)

It is interesting to note that the measurement error ΔX_f in position and the measurement error $\delta p_{\psi f}$ in momentum obey an uncertainty relation

$$\Delta X_f\delta p_f\cong(\pi/\sqrt{3})\hbar>\hbar/2. \tag{17}$$

This result has sometimes been confused with the intrinsic state uncertainty relation (3).

E. Joint-measurement uncertainty relation

The total uncertainties for the position and momentum measurements have a product

$$\Delta X_{\psi f}\delta p_{\psi f}\cong\sqrt{(\Delta x_\psi^2+w^2/12)(\Delta p_\psi^2+(2\pi\hbar)^2/w^2)}, \tag{18a}$$

which follows from Eqs. (12b) and (15). We may vary the slit width w to minimize this uncertainty product. The minimum value is found to be

$$\Delta X_{\psi f}\delta p_{\psi f}\cong\Delta x_\psi\Delta p_\psi+2\pi\hbar/2\sqrt{3}, \tag{18b}$$

showing the added uncertainty arising from the joint measurement. Invoking the state uncertainty relation (3) we arrive at

$$\Delta X_{\psi f}\delta p_{\psi f}\geq\frac{\hbar}{2}\left(1+\frac{2\pi}{\sqrt{3}}\right)\cong\frac{\hbar}{2}(4.6), \tag{19}$$

the minimum value of which occurs for $w=\sqrt{2\pi\sqrt{3}}\Delta x_\psi$. The product in Eq. (19) is about four times larger than the intrinsic state uncertainty $\hbar/2$ in Eq. (3). This is an important illustrative result of this paper, as it shows that when attempting to make joint measurements on two noncommuting variables the measurements are imprecise, and this leads unavoidably to additional uncertainties beyond those implied by the usual (state) uncertainty principle (3).

Actually, the specific result (19) does not represent a general lower bound for joint measurement, as it was derived only for a particular example and was not derived quantitatively, due to the approximations in Eqs. (12b) and (15). This treatment was instigated by the divergence of the momentum variance caused by diffraction from the sharp slit edges. There are two attitudes that have been taken in the literature with regard to this problem: Hilgevoord and Uffink have argued that generally the variance is not a good measure of uncertainty.[9] They have instead derived an uncertainty relation based on the widths of the distributions. They define the width as that range of x containing some specified fraction (near unity) of the total probability. On the other hand, Braginsky and Khalili (among others) have adopted the attitude that the problem is not in the use of the variance, but in the assumption that measurements can be made with sharp boundaries between possible measured values (sharp slit in our case), and with nonoverlapping measurement channels.[10]

In their approach the channels do not have sharp edges and are overlapping (nonorthogonal). In Sec. III I will adopt this latter approach and derive again the joint-measurement uncertainty principle in a more general context. This results in a smaller uncertainty product than in Eq. (18), but, as we will see below, never as small as the intrinsic state uncertainty in Eq. (3).

III. UNCERTAINTY RELATIONS WITH JOINT, NONORTHOGONAL MEASUREMENTS

A. Generalized measurements

In a modern version of the measurement model, which is a generalization of von Neumann's projection model, the conditional probability $f^2(X|x)$ in Eq. (8b) need not be a sharp-edged top-hat function.[6,10,18–25] If we instead allow this function to be smoothed out, the divergence of the variance of transverse momentum is avoided. In some sense this is more realistic than assuming perfectly sharp edges. Rather than imagining a sharp slit as the measurement device, consider the scattering of a light beam from the particle as a way to measure its position. This is Heisenberg's famous microscope argument that he used in his early discussions of the uncertainty principle.[8] The optical resolution of the microscope limits the precision with which the particle's position can be inferred. Rayleigh's resolving criterion leads to an overlapping of the different resolvable "spots" or nonorthogonal resolution functions.

For example, the conditional probability that, given the particle is perfectly localized at x, the measurement will yield X might be given by a Gaussian function,

$$f^2(X|x)=(\pi 2\sigma^2)^{-1/2}\exp[-(X-x)^2/2\sigma^2], \tag{20}$$

where σ is the resolution spot size. Each resolution function is centered on a different value of X, which is continuously distributed. This function is represented pictorially in Fig. 2(b). The conditional probability is normalized,

$$\int_{-\infty}^{\infty} f^2(X|x)dX=1, \tag{21}$$

since every measurement must yield some result. The X-measuring apparatus is fully characterized by $f^2(X|x)$.

It greatly simplifies the analysis to assume that the resolution function depends only on the difference variable $X-x$ and that it is symmetric around its maximum value which occurs at $X=x$, as, e.g., in Eq. (20). I will thus denote $f^2(X|x)$ by $f^2(X-x)$. This is possible in the present case because position and momentum are continuous variables with no upper or lower bounds, and the commutator between the variables is a c number.[10] There are pairs of conjugate variables other than position and momentum (e.g., photon number and optical phase) for which this simple model is not applicable.

The probability to measure a position value X is then given by

$$P_{\psi f}(X)=\int_{-\infty}^{\infty} f^2(X-x)|\psi(x)|^2dx. \tag{22}$$

Because of the symmetry of $f^2(X-x)$, the mean value $\bar{X}$ of X equals the mean $\bar{x}_\psi$ of the state distribution—that is, the measurement is unbiased. The measurement error (resolution) is the variance of the conditional probability,

$$\Delta X_f^2=\int_{-\infty}^{\infty}(X-x)^2f^2(X-x)dX, \tag{23}$$

and is independent of the value of x. Because of the convolution form of Eq. (22), the standard deviation [square root of Eq. (11)] of the position measurements is given by

$$\Delta X_{\psi f}=\sqrt{\Delta x_\psi^2+\Delta X_f^2}. \tag{24}$$

Equation (24) holds for arbitrary forms of the functions under the integral in Eq. (22), and is an exact result, in contrast to the similar expression (12b).

In a generalization of Eq. (14), the wave function $\psi_{fX}(x)$ after the position measurement yields a particular X is given by[6,24,25]

$$\psi_{fX}(x)=\frac{f(X-x)\psi(x)}{\sqrt{\int f^2(X-x)|\psi(x)|^2dx}}, \tag{25}$$

where $f(X-x)$ is the real, positive square root of $f^2(X-x)$. In contrast to Eq. (14), this result is not a projection of the incident state, because $f^2(X-x)$ is not proportional to $f(X-x)$. This means that immediately successive position measurements will not necessarily give the same results. This fact is actually a strength of the present model, not a weakness, even though it differs from the standard von Neumann model.[10] It is a strength because, for example, it shows that successive position measurements will yield an increasingly better determination of the position, through "signal averaging," a fact known to any experimenter.

Related generalizations of von Neumann's projection model have been advocated by several authors, in rather formal discussions.[18–22] These have been reviewed by Busch *et al.*,[23] Braunstein *et al.*,[24] and Gardiner.[25] In these models, von Neumann's projection operators are replaced by "effect-valued (or positive-operator-valued) measures." These models are not simply postulated, but rather are derived from standard quantum treatments of model interactions between object and measuring apparatus. The results derived in the present paper are consistent with these more formal treatments.

B. Joint measurements and uncertainty relations

The transverse momentum distribution, after the position measurement, is given by $|\tilde{\psi}_{fX}(p)|^2$ where $\tilde{\psi}_{fX}(p)$ is the Fourier transform of $\psi_{fX}(x)$. Again we assume precise measurement of momentum of the particle after the position measurement takes place. Because $|\tilde{\psi}_{fX}(P)|^2$ is the conditional probability to observe momentum P given a position measurement yielding X, we can define a joint probability for position and momentum measurements as

$$\begin{aligned}P_{\psi f}(X,P)&=|\tilde{\psi}_{fX}(P)|^2P_{\psi f}(X)\\&=(2\pi\hbar)^{-1}\left|\int_{-\infty}^{\infty}f(X-x)\psi(x)\exp(-iPx/\hbar)dx\right|^2.\end{aligned} \tag{26}$$

This is a central result of the theory of joint measurement.[4,6,10]

The probability to observe momentum P independently of the observed X value is

$$\bar{P}_{\psi f}(P)=\int_{-\infty}^{\infty}P_{\psi f}(X,P)dX. \tag{27}$$

This can be rewritten as

$$\bar{P}_{\psi f}(P)=\int_{-\infty}^{\infty}|\tilde{f}(P-p)|^2|\tilde{\psi}(p)|^2dp, \tag{28}$$

where $\tilde{f}(p)$ is the Fourier transform [defined by Eq. (2)] of $f(x)$. Note that this probability is equal to a convolution of the momentum distribution of the incident state with the momentum resolution function $|\tilde{f}(P-p)|^2$. This result is appealing because it has the same form as the position distribution in Eq. (22). So it appears that although position is measured first in the example used, actually position and momentum measurements are on equal footing in this formalism. The same results would be obtained if momentum were measured first, followed by position measurement. In this case one can use the term "simultaneous measurement," which is actually a more commonly used term than is joint measurement.

The measurement error variance for momentum is

$$\Delta P_f^2=\int_{-\infty}^{\infty}(P-p)^2|\tilde{f}(P-p)|^2dP, \tag{29}$$

and is a property only of the measuring apparatus. The standard deviation ΔP_f of $|\tilde{f}(P-p)|^2$ and the standard deviation ΔX_f of $f^2(X-x)$ satisfy the relation

$$\Delta X_f\Delta P_f \geqslant \hbar/2, \tag{30}$$

as a consequence of the properties of the Fourier transform, the relation being an equality only for Gaussian functions. [The $\hbar$ enters as a consequence of the definition of the transform in Eq. (2).] This may be called the measurement uncertainty (or error) relation. It has the familiar interpretation that the smaller the measurement error in position is, the larger is the perturbation of momentum and consequent momentum measurement error. But note that this is not the same as the intrinsic state uncertainty relation (3). Equation (30) is a property of the measurement apparatus, not the state of the particle.

Finally we may formulate the uncertainty relation for joint measurement. From the distribution (28) we can find the standard deviation of momentum measurements (which is now finite),

$$\Delta P_{\psi f}=\sqrt{\Delta p_\psi^2+\Delta P_f^2}, \tag{31}$$

where Δp_ψ is again the intrinsic state uncertainty. The uncertainty product is

$$\Delta X_{\psi f}\Delta P_{\psi f}=\sqrt{(\Delta x_\psi^2+\Delta X_f^2)(\Delta p_\psi^2+\Delta P_f^2)}. \tag{32}$$

This can be rewritten solely in terms of Δx_ψ and ΔX_f by using

$$\begin{aligned}\Delta p_\psi&=\alpha_\psi\hbar/2\Delta x_\psi,\\ \Delta P_f&=\alpha_f\hbar/2\Delta X_f,\end{aligned} \tag{33}$$

where α_ψ and α_f satisfy $\alpha_\psi \geqslant 1$, $\alpha_f \geqslant 1$, and are measures of how close Eqs. (3) and (30) come to being equalities. If $\alpha_\psi=1$ we say that the initial state is a minimum-uncertainty state. If $\alpha_f=1$ we will say that the measurement is a minimum-error measurement. [Note that in the earlier example in Sec. II α_f was fixed at the value $2\pi\sqrt{3}$, as seen from Eq. (17).] The uncertainty product (32) can be minimized by adjusting the variance of the position resolution function; the minimum occurs for $\Delta X_f^2=(\alpha_f/\alpha_\psi)\Delta x_\psi^2$, and gives the extended uncertainty relation,

$$\Delta X_{\psi f}\Delta P_{\psi f} \geqslant (\alpha_\psi+\alpha_f)\hbar/2. \tag{34}$$

The optimum case occurs for $\alpha_\psi=\alpha_f=1$, which can occur only if both the state and the filter are Gaussian in form. In this case the uncertainty relation for joint measurement becomes

$$\Delta X_{\psi f}\Delta P_{\psi f} \geqslant \hbar. \tag{35}$$

This uncertainty product implies a lower bound that is at least twice as large as the intrinsic state uncertainty product (3) or the measurement error product (30). It shows that some information can be obtained about the distribution of both position and momentum by observing imprecisely both of these variables for each particle in the ensemble. The price that is paid for this is the introduction of error beyond that contained in the intrinsic state uncertainty.

An interesting special case occurs if $\alpha_f \ll \alpha_\psi$, that is, the measurement is much closer to being optimal than is the state. Then Eq. (34) shows that joint measurement of x and p can be accomplished without adding any significant uncertainty beyond that contained in the state.

IV. CONNECTION TO ARTHURS–KELLY THEORY OF JOINT MEASUREMENT

The convolution form of the probability functions (22) and (28) can be used to postulate formally the existence of new quantum variables, A and B, which are independent of the particle's variables x and p, and which add to them during the course of a "measurement" interaction. In operator form

$$\hat{X}=\hat{x}+\hat{A}, \quad \hat{P}=\hat{p}+\hat{B}. \tag{36}$$

Recall that in the theory of probability the distribution of a variable that is the sum of two independent variables is given by the convolution of their respective distributions. The functions f^2 and $|\tilde{f}|^2$ serve as distributions for $\hat{A}$ and $\hat{B}$, respectively. The sum variables, $\hat{X}$ and $\hat{P}$ play the role of meters which can be read out after interacting with the particle. The new variables, $\hat{A}$ and $\hat{B}$ play the role of meter noise operators.

Equation (36) is postulated in the Arthurs–Kelly theory of joint measurement,[4] as reviewed by Yamamoto and Haus[5] and by Stenholm.[6] The basic idea is to chose $\hat{A}$ and $\hat{B}$ such that $\hat{X}$ and $\hat{P}$ commute, and so can be read separately without introducing any noise beyond that introduced by $\hat{A}$ and $\hat{B}$. This is achieved only if $[\hat{A},\hat{B}]=-[\hat{x},\hat{p}]=-i\hbar$, which also implies that the standard deviations of the noise operators obey

$$\Delta A\Delta B \geqslant \hbar/2. \tag{37}$$

If the particle operators and noise operators are uncorrelated, then the variances add and the uncertainty product for $\hat{X}$ and $\hat{P}$ is

$$\Delta X\Delta P=\sqrt{(\Delta x^2+\Delta A^2)(\Delta p^2+\Delta B^2)}. \tag{38}$$

This is precisely of the form of Eq. (32) and so implies the same bounds given by Eqs. (34) and (35). This shows that the measurement theory used in Sec. III is consistent with that of Arthurs and Kelly. This further helps establish that all results discussed here can be understood in terms of standard quantum mechanics.

V. CONNECTION TO THE WIGNER FUNCTION

Finally I would point out that the general result (26) for the joint measurement of x and p can be expressed in terms of the Wigner functions representing the state and the measurement device. For a wave function $\psi(x)$ the associated Wigner function is defined as[26]

$$W_{\psi}(x,p)=(\pi\hbar)^{-1}\int_{-\infty}^{\infty}\psi(x+x')\psi^{*}(x-x')\times\exp(-2ipx'/\hbar)dx'. \quad (39)$$

There is an analogously defined Wigner function $W_f(x,p)$ associated with the resolution function $f(x)$. It is easy to show that the joint probability (26) for measurement of x and p can be expressed as a convolution of these two Wigner functions

$$P_{\psi f}(X,P)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}W_f(X-x,p-P)W_{\psi}(x,p)dxdp. \quad (40)$$

This result is of the same form as discussed in Lai and Haus,[11] who treat the joint measurement of the two quadrature components (in-phase and 90°-out-of-phase electric-field amplitudes) of a light beam. The connection between Eqs. (26) and (40) has been pointed out by Stenholm.[6] It has also been discussed by Brenner and Wodkiewicz in a different context: filtering in the time-frequency domain.[27] Equation (40) shows that the Wigner function, which is not a true joint probability (since it can take on negative values), is converted to the true joint probability after convolution with the measurement resolution function.

VI. EXPERIMENTS

Joint measurements of the kind discussed here have been carried out in the case of single-mode optical fields.[28] The two conjugate variables are the real (cos) and imaginary (sin) parts of the oscillating electric field, called quadrature amplitudes. The measurement is made by beam splitting the field with a 50%/50% beam splitter and then using coherent detection techniques (balanced homodyne, or interferometric, detection) to measure the real part of one beam and the imaginary part of the other beam. The additional noise comes into the signal field by virtue of the random nature of optical beam splitting: each photon has a nonzero chance of reflecting or transmitting through the beam splitter. On the other hand, when no beam splitter is used, and the whole beam is incident on the coherent detector, the real or imaginary parts (but not both simultaneously) can be measured with no added noise. In this case, the Wigner function (39) of the signal field can be determined from the measurements, as demonstrated experimentally by Smithey *et al.*[29]

VII. CONCLUSIONS

Using a simple, familiar example I have verified that, according to standard quantum mechanics, if joint (although necessarily imprecise) measurements are made on *both* position and conjugate momentum for each member of an ensemble of identically prepared systems, the product of standard deviations for the measurement outcomes is at least twice as large as the lower bound implied by the usual uncertainty principle. The result is stated in Eq. (35), which was previously known.[4–6] I have provided a somewhat more informative version of this result in Eq. (34), in which it can be seen how the contributions from the intrinsic state indeterminacy and the measurement error combine to give the overall uncertainty product. For example, if the measurement is optimal but the state has uncertainty product much greater than that of a minimum-uncertainty state, then Eq. (34) shows that joint measurement of a pair of noncommuting variables can be accomplished without adding any significant extra uncertainty beyond that contained in the state. Connections have been made between these results and those of other authors,[4,10,11] showing them all to be alternative methods by which to treat the same physics.

This pedagogical discussion helps to clarify the different roles of the initial state and of the measurement in the various statements of the uncertainty principle. Heisenberg's original argument[7] (as reviewed in Ref. 9) was that if one measures x with resolution (error) ΔX_f, then one can predict p in a subsequent measurement with an accuracy (error) of order ΔP_f, where $\Delta X_f \Delta P_f \gtrsim h$. This is essentially Eq. (30), and by comparing to Eq. (32), we can see that Heisenberg's argument does not deal with the incident state uncertainties Δx_{ψ} and Δp_{ψ}, but only with the measurement uncertainties. Of course there is a close connection between these concepts, since measuring x serves as a state preparation, due to the state reduction, or backaction, represented in Eq. (25). But note that Heisenberg's formulation does not involve varying the position of the slit in array 1 in order to infer information about the incident state $\psi(x)$. It is only concerned with state reduction (and backaction) at the slit and subsequent measurement of momentum. A related discussion is given by Martens and de Muynck.[30]

Heisenberg's original formulation, using the microscope gedanken experiment as motivation, is often cited in textbooks as the essence of the uncertainty principle. On the other hand, when a quantitative derivation of the uncertainty relation is desired, most textbooks (including Heisenberg's,[8] but excepting Ballentine's[31]) adopt Kennard's approach[13] using the variance of the intrinsic state uncertainties, and use the Fourier properties of the wave function (or equivalently, operator algebra)[32] to derive the relation (3). But since this involves only the intrinsic state uncertainties, it is not actually germane to Heisenberg's original ideas. Thus, one argument considered only measurement uncertainties [Eq. (30)], while the other argument considered only initial state uncertainties [Eq. (3)]. The total uncertainty for joint measurement is a combination of both of these uncertainties [Eq. (32)]. By considering both the state indeterminacy and the measurement error in a unified discussion, as discussed here, the physics becomes clearer.

APPENDIX

Equation (15) for the estimated full width of the transverse momentum (p) distribution after passing the slit can be obtained as follows. First consider the case in which the width of the wave function in x is much broader than the width of the slit. Then, according to Eq. (14), the wave function after the slit is proportional to the top-hat function $h(X_J-x)$. Therefore the momentum distribution is proportional to

$$|\tilde{h}(p)|^2 \propto \left|\frac{\sin(pw/2\hbar)}{p}\right|^2. \quad (A1)$$

This distribution has a full width at half-maximum equal to about $\delta p_f = 2\pi\hbar/w$. This verifies Eq. (15) in the limit $\Delta p_\psi \ll \delta p_f$. The opposite limit simply corresponds to no slit being present. The root-mean-square form of Eq. (15) is then chosen arbitrarily to interpolate smoothly between these two limits.

ACKNOWLEDGMENTS

I wish to thank M. Beck, J. Cooper, and D. Soper for critical readings of this manuscript. This research was supported by the National Science Foundation under Grant No. PHY-924779.

[a]raymer@oregon.uoregon.edu

[1]For review of quadrature-amplitude-squeezed states of light see K. Zaheer and M. S. Zubairy, "Squeezed light," in *Advances in Atomic, Molecular, and Optical Physics*, edited by D. Bates and B. Bederson (Academic, Boston, 1991), Vol. 28, pp. 143–236.

[2]M. Sigel and J. Mlynek, "Atom optics," Phys. World **6**, 36–42 (1993).

[3]*The Detection of Gravitational Waves*, edited by D. G. Blair (Cambridge, New York, 1991).

[4]E. Arthurs and J. L. Kelly, "On the simultaneous measurement of a pair of conjugate observables," Jr., Bell Syst. Tech. J. **44**, 725–729 (1965).

[5]Y. Yamamoto and H. A. Haus, "Preparation, measurement and information capacity of optical quantum states," Rev. Mod. Phys. **58**, 1001–1020 (1986).

[6]S. Stenholm, "Simultaneous measurement of conjugate variables," Ann. Phys. **218**, 233–254 (1992).

[7]The first paper on the uncertainty principle is: W. Heisenberg, "Über den anschaulichen Inhalt der quantentheoretischen kinematik und mechanik," Z. Phys. **43**, 172–198 (1927); English translation can be found in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. Zurek (Princeton University Press, Princeton, NJ, 1983), pp. 62–84.

[8]W. Heisenberg, *The Physical Principles of the Quantum Theory*, English translation (Dover, New York, 1930).

[9]J. Hilgevoord and J. Uffink, "A new view on the uncertainty principle," in *Sixty-Two Years of Uncertainty*, edited by A. I. Miller (Plenum, New York, 1990), pp. 121–137.

[10]V. B. Braginsky and F. Ya. Khalili, *Quantum Measurement* (Cambridge University Press, Cambridge, MA, 1992), especially Chaps. I, II, and V.

[11]Y. Lai and H. A. Haus, "Characteristic functions and quantum measurements of optical observables," Quantum Opt. **1**, 99–115 (1989).

[12]C. Cohen-Tannoudji, B. Diu, and F. Laloe, *Quantum Mechanics* (Wiley, New York, 1977), pp. 28 and 231.

[13]E. H. Kennard, "Zur quantenmechanik einfacher bewegungstypen," Z. Phys. **44**, 326–352 (1927).

[14]E. T. Jaynes, "Clearing up mysteries: The original goal," in *Maximum Entropy and Bayesian Methods in Applied Statistics*, edited by J. Skilling (Kluwer Academic, Dordrecht, 1989), pp. 1–27; for elementary discussion of Bayes' theorem see P. Z. Peebles, *Probability, Random Variables, and Random Signal Principles* (McGraw–Hill, New York, 1980).

[15]J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, English translation (Princeton University Press, Princeton, NJ, 1955), especially Chaps. V and VI.

[16]See also Ref. 12, p. 265.

[17]Equation (14) can also be obtained from a more formal argument, which will be important only in the context of the generalized measurement model discussed in Sec. III. If the particle is found in the window at X_J, it can equally well be inferred to have been located at any particular position X within the window (since all values of X within the window are equally likely). Then the state becomes

$$\psi_{fX}(x) = [f(X|x)\psi(x)]/\sqrt{\int f^2(X|x)|\psi(x)|^2\,dx},$$

where $f(X|x)$ is the square root of $f^2(X|x)$ defined in Eq. (9). Because X is contained within the window at X_J, then $f^2(X|x)$ becomes $f_J^2(X|x) = w^{-1}h(X_J - x)$, and this result reduces to Eq. (14).

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

[19]E. B. Davies, *Quantum Theory of Open Systems* (Academic, London, 1976).

[20]G. Ludwig, *Foundations of Quantum Mechanics* (Springer, Berlin, 1983), Vol. 1.

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

[22]A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).

[23]P. Busch, P. J. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement* (Springer, Berlin, 1991).

[24]S. L. Braunstein, C. M. Caves, and G. J. Milburn, "Interpretation for a positive P representation," Phys. Rev. A **43**, 1153–1159 (1991), Sec. III; see also C. M. Caves, "Quantum mechanics of measurements distributed in time: a path integral formulation," Phys. Rev. D **33**, 1643–1665 (1986), Secs. II A and II B.

[25]C. W. Gardiner, *Quantum Noise* (Springer, Berlin, 1991), Sec. 2.2.

[26]E. P. Wigner, "On the quantum correction for thermodynamic equilibrium," Phys. Rev. **40**, 749–759 (1932); M. Hillery, R. F. O'Connell, M. O. Scully, and E. P. Wigner, "Distribution functions in physics: Fundamentas," Phys. Rep. **106**, 121–167 (1984).

[27]K.-H. Brenner and K. Wodkiewicz, "The time-dependent physical spectrum of light and the Wigner distribution function," Opt. Commun. **43**, 103–106 (1982).

[28]N. G. Walker and J. E. Carroll, "Multiport homodyne detection near the quantum noise limit," Opt. Quantum Electron. **18**, 355–363, (1986); N. G. Walker, "Quantum theory of multiport optical homodyning," J. Mod. Opt. **34**, 16–60 (1987).

[29]D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, "Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum," Phys. Rev. Lett. **70**, 1244–1247 (1993); D. T. Smithey, M. Beck, J. Cooper, M. G. Raymer, and A. Faridani, "Complete experimental characterization of the quantum state of a light mode via the Wigner function and the density matrix: application to quantum phase distributions of vacuum and squeezed-vacuum states," Phys. Scr. **T48**, 35–44 (1993).

[30]H. Martens and W. de Muynck, "Towards a new uncertainty principle: Quantum measurement noise," Phys. Lett. **157**, 441–448 (1991); "Disturbance, conservation laws and the uncertainty principle," J. Math Phys. **25**, 4887–4901 (1992).

[31]L. E. Ballentine, *Quantum Mechanics* (Prentice–Hall, Englewood Cliffs, NJ, 1990).

[32]H. P. Robertson, "The uncertainty principle," Phys. Rev. **34**, 163–164 (1929).

GRAVITATION AND CONSCIOUSNESS

Early in the book, Todd Siler compares neurons with stars and brains with galaxies. Just as neurons talk to one another with electrochemical pulses, he proposes, stars communicate in the language of gravity. Or as he puts it: "Gravitation is to galaxies what consciousness is to human beings."

George Johnson, in a review of Todd Siler, *Breaking the Mind Barrier*, The New York Times Book Review, December 30, 1990, p. 7.