International Journal of Theoretical Physics, Vol. 37, No. 10, 1998

Error Principle

D. M. Appleby[1]

Received July 2, 1998

The problem of characterizing the accuracy of and disturbance caused by a joint measurement of position and momentum is investigated. In a previous paper the problem was discussed in the context of the unbiased measurements considered by Arthurs and Kelly. It is now shown that suitably modified versions of these results hold for a much larger class of simultaneous measurements. The approach is a development of that adopted by Braginsky and Khalili in the case of a single measurement of position only. A distinction is made between the errors of retrodiction and the errors of prediction. Two error–error relationships and four error–disturbance relationships are derived, supplementing the uncertainty principle usually so-called. In the general case it is necessary to take into account the range of the measuring apparatus. Both the ideal case of an instrument having infinite range and the case of a real instrument for which the range is finite are discussed.

1. INTRODUCTION

In Appleby (1998a) we derived a number of inequalities relating the accuracy of and disturbance caused by a simultaneous measurement of position and momentum. However, our discussion was limited to the special case of unbiased measurement processes, for which the systematic errors are zero. Our purpose in the following is to generalize the results we obtained earlier and show that analogous relations can be derived for a very much larger class of measurement processes.

Heisenberg's (1927, 1930) formulation of the uncertainty principle was one of the key steps in the development of quantum mechanics. Nevertheless, over 70 years after the publication of his original paper, there remain a number of obscurities regarding its interpretation (Hilgevoord and Uffink,

[1] Department of Physics, Queen Mary and Westfield College, London E1 4NS, U.K.; e-mail: D.M.Appleby@qmw.ac.uk.

0020-7748/98/1000-255 7$15.00/0

1990). In contemporary discussions the uncertainty principle is usually identified with the statement

$$\Delta x \Delta p \geq \frac{\hbar}{2} \tag{1}$$

where Δx, Δp are the standard deviations

$$\begin{aligned} \Delta x &= (\langle\psi|\hat{x}^2|\psi\rangle - \langle\psi|\hat{x}|\psi\rangle^2)^{1/2} \\ \Delta p &= (\langle\psi|\hat{p}^2|\psi\rangle - \langle\psi|\hat{p}|\psi\rangle^2)^{1/2} \end{aligned} \tag{2}$$

In his original paper Heisenberg suggested that the quantities Δx, Δp appearing in (1) may be interpreted as experimental errors, and that the uncertainty principle represents a fundamental constraint on the accuracy achievable in a simultaneous measurement of position and momentum. At least, that is what he has often been taken to have suggested (Heisenberg's own phraseology is somewhat ambiguous). In the words of Bohm (1951), "If a measurement of position is made with accuracy Δx, and if a measurement of momentum is made *simultaneously* with accuracy Δp, then the product of the two errors can never be smaller than a number of order $\hbar$."

Is this is a legitimate interpretation of (1)? The question has been discussed by Ballentine (1970), Prugovečki (1973, 1984), Busch (1985), Wódkiewicz (1987), Hilgevoord and Uffink (1990), Raymer (1994), de Muynck *et al.* (1994), and Busch *et al.* (1995). The consensus seems to be that the quantities Δx, Δp defined in (2) cannot be regarded as experimental errors because they are intrinsic properties of the isolated system. An experimental error, by contrast, should depend, not only on the state of the system, but also on the state of the apparatus and the nature of the measurement interaction. Hilgevoord and Uffink (1990) have further remarked that in Heisenberg's microscope argument it is only the position of the particle which is measured. Although it is true that Heisenberg alludes to the possibility of performing simultaneous measurements of position *and* momentum, such measurements form no part of his actual argument.

It follows from all this that the statement of Bohm's just quoted cannot be identified with the uncertainty principle usually so-called. Rather, it represents (if true) an independent physical principle: the error principle, as it might be called.

The problem we now face is that although the error principle as stated by Bohm is intuitively quite plausible, it cannot be regarded as rigorously established. In order to establish it two things are necessary. In the first place, we need to define precisely what is meant by the accuracy of a simultaneous measurement process. In the second place, we need to derive a bound on the accuracy, starting from the fundamental principles of quantum mechanics. The problem is of some interest, in view of the importance that simultaneous measurements now have in the field of quantum optics (Arthurs and Kelly, 1965; Arthurs and Goodman, 1988; Braunstein *et al.*, 1991; Stenholm, 1992; Leonhardt and Paul, 1993a,b, 1995b; Törma *et al.*, 1995; Busch *et al.*, 1995; Power *et al.*, 1997; Leonhardt, 1997).

An approach to the problem which has attracted a good deal of attention over the years is the one based on positive-operator-valued measures and the concept of a "unsharp" observable (Prugovečki, 1973, 1976a,b, 1978, 1984; Davies, 1976; Holevo, 1982; Busch, 1985; Busch and Lahti, 1989; Martens and de Muynck, 1990a,b, 1992; deMuynck *et al.*, 1994; Ban, 1997). For a recent review see Busch *et al.* (1995). This approach has recently been criticized by Uffink (1994).

In Appleby (1998a) we adopted a rather different approach. We began with Braginsky and Khalili's (1992) analysis of single measurements of x or p by themselves and extended it to a class of simultaneous measurement processes, namely, the class of unbiased measurement processes, for which the systematic errors are all zero. Our analysis depended on making a distinction between the retrodictive and predictive (or determinative and preparative) aspects of a measurement (Margenau, 1958; Prugovečki, 1973, 1976; Hilgevoord and Uffink, 1990; Busch and Lahti, 1989). We accordingly defined two different kinds of error: the errors of retrodiction, $\Delta_{ei}x$ and $\Delta_{ei}p$, describing the accuracy with which the result of the measurement reflects the initial state of the system; and the errors of prediction, $\Delta_{ef}x$ and $\Delta_{ef}p$, describing the accuracy with which the result of the measurement reflects the final state of the system. Corresponding to these two kinds of error we derived two inequalities: a retrodictive error relationship

$$\Delta_{ei}x \, \Delta_{ei}p \geq \frac{\hbar}{2} \tag{3}$$

and a predictive error relationship

$$\Delta_{ef}x \, \Delta_{ef}p \geq \frac{\hbar}{2} \tag{4}$$

Relations (3) and (4) jointly comprise a precise statement of the semiintuitive error principle discussed above.

Following Braginsky and Khalili, we also defined two quantities $\Delta_d x$, $\Delta_d p$ describing the disturbance of the system by the measurement, and we derived the four error–disturbance relationships

$$\Delta_{\text{ei}} x \, \Delta_{\text{d}} p \geq \frac{\hbar}{2}, \qquad \Delta_{\text{ef}} x \, \Delta_{\text{d}} p \geq \frac{\hbar}{2} \tag{5}$$

$$\Delta_{\text{ei}} p \, \Delta_{\text{d}} x \geq \frac{\hbar}{2}, \qquad \Delta_{\text{ef}} p \, \Delta_{\text{d}} x \geq \frac{\hbar}{2}$$

These relationships provide a precise statement of the principle that a decrease in the error of the measurement of one observable can only be achieved at the cost of a corresponding increase in the disturbance of the canonically conjugate observable.

The relationships above, together with (1), comprise a total of seven inequalities, all of which are needed if one wants to capture the full intuitive content of Heisenberg's (1927) original paper.

Arthurs and Kelly (1965) have shown that in the case of a retrodictively unbiased joint measurement process (i.e., a process for which the final state expectation values of the pointer positions coincide with the initial state expectation values of the position and momentum), one has

$$\Delta\mu_{\text{X}} \, \Delta\mu_{\text{P}} \geq \hbar \tag{6}$$

where the quantities on the right-hand side are the final state uncertainties for the pointer positions μ_{X} and μ_{P} (also see Arthurs and Goodman, 1988; Wódkiewicz, 1987; Raymer, 1994; Leonhardt and Paul, 1995a). In Appleby (1998a) we showed that the Arthurs–Kelly relationship can be deduced from the retrodictive error relationship.

The arguments given in Appleby (1998a) only serve to establish the above inequalities for a limited class of measurement processes. That is, we only proved (3) on the assumption that the measurement is retrodictively unbiased, and (5) on the still more restrictive assumption that the measurement is both retrodictively and predictively unbiased. Our purpose in the following is to show that with a suitable modification of the definitions, these relationships continue to hold for a very much larger class of measurement processes.

2. SIMULTANEOUS MEASUREMENT PROCESSES

We begin by characterizing the class of measurement processes which we are going to discuss.

Consider a system with state space $\mathscr{H}_{\text{sy}}$ interacting with an apparatus with state space $\mathscr{H}_{\text{ap}}$. The system is assumed to have one degree of freedom, with position $\hat{x}$ and momentum $\hat{p}$ satisfying the commutation relationship

$$[\hat{x}, \hat{p}] = i\hbar \tag{7}$$

The apparatus is assumed to be characterized by two pointer observables $\hat{\mu}_{\text{X}}$

(measuring the position of the system) and $\hat{\mu}_P$ (measuring the momentum of the system), together with n other observables $\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n$. These $n + 2$ operators constitute a complete set of commuting observables describing the state of the apparatus. They also commute with the system observables $\hat{x}, \hat{p}$.

It is assumed that the system + apparatus is initially in a product state of the form $|\psi \otimes \phi_{ap}\rangle$, where $|\psi\rangle \in \mathcal{H}_{sy}$ is the initial state of the system and $|\phi_{ap}\rangle \in \mathcal{H}_{ap}$ is the initial state of the apparatus. The unitary evolution operator describing the measurement interaction will be denoted $\hat{U}$. The final state of the system + apparatus is $\hat{U}\,|\psi \otimes \phi_{ap}\rangle$. The probability distribution of the measured values is

$$\rho(\mu_X, \mu_P) = dx\, dy_1 \ldots dy_n |\langle x, \mu_X, \mu_P, y_1, \ldots, y_n|\hat{U}\,|\psi \otimes \phi_{ap}\rangle|^2$$

In Appleby (1998a) we assumed that the measurement process was unbiased, so that

$$\langle\psi\otimes\phi_{ap}|\hat{U}^\dagger\,\hat{\mu}_X\,\hat{U}\,|\psi\otimes\phi_{ap}\rangle = \langle\psi\otimes\phi_{ap}|\hat{U}^\dagger\,\hat{x}\hat{U}\,|\psi\otimes\phi_{ap}\rangle = \langle\psi\otimes\phi_{ap}|\hat{x}\,|\psi\otimes\phi_{ap}\rangle$$

and

$$\langle\psi\otimes\phi_{ap}|\hat{U}^\dagger\,\hat{\mu}_P\hat{U}\,|\psi\otimes\phi_{ap}\rangle = \langle\psi\otimes\phi_{ap}|\hat{U}^\dagger\hat{p}\hat{U}\,|\psi\otimes\phi_{ap}\rangle = \langle\psi\otimes\phi_{ap}|\,\hat{p}\,|\psi\otimes\phi_{ap}\rangle$$

We make no such assumption here.

It may also be worth noting that we do not assume the existence of momenta canonically conjugate to the pointer observables (as is the case in the Arthurs–Kelly process, for example [Arthurs and Kelly, 1965; Busch, 1985; Braunstein *et al.*, 1991; Stenholm, 1992; Power *et al.*, 1997; Leonhardt, 1997; Appleby, 1998a,b]). In particular, we make no assumptions regarding the spectra of the pointer observables.

3. DEFINITION OF THE ERRORS AND DISTURBANCES

Let $\mathcal{O}$ be any of the Schrödinger-picture operators $\hat{x}, \hat{p}, \hat{\mu}_X, \hat{\mu}_P$. Let $\mathcal{O}_i = \mathcal{O}$ be the corresponding Heisenberg-picture operator at the instant the measurement interaction begins, and let $\mathcal{O}_f = \hat{U}^\dagger \mathcal{O} \hat{U}$ be the Heisenberg-picture operator at the instant the interaction finishes. Define the retrodictive error operators

$$\hat{\varepsilon}_{Xi} = \hat{\mu}_{Xf} - \hat{x}_i, \qquad \hat{\varepsilon}_{Pi} = \hat{\mu}_{Pf} - \hat{p}_i \tag{8}$$

the predictive error operators

$$\hat{\varepsilon}_{Xf} = \hat{\mu}_{Xf} - \hat{x}_f, \qquad \hat{\varepsilon}_{Pf} = \hat{\mu}_{Pf} - \hat{p}_f \tag{9}$$

and the disturbance operators

$$\hat{\delta}_{\mathrm{X}} = \hat{x}_{\mathrm{f}} - \hat{x}_{\mathrm{i}}, \qquad \hat{\delta}_{\mathrm{P}} = \hat{p}_{\mathrm{f}} - \hat{p}_{\mathrm{i}} \tag{10}$$

Let $\mathcal{S}$ be the unit sphere in the system state space $\mathcal{H}_{\mathrm{sy}}$. We then define the maximal rms errors of retrodiction

$$\begin{aligned} \Delta_{\mathrm{ei}} x &= \sup_{|\psi\rangle \in \mathcal{S}} (\langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}_{\mathrm{Xi}}^2 | \psi \otimes \phi_{\mathrm{ap}} \rangle)^{1/2} \\ \Delta_{\mathrm{ei}} p &= \sup_{|\psi\rangle \in \mathcal{S}} (\langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}_{\mathrm{Pi}}^2 | \psi \otimes \phi_{\mathrm{ap}} \rangle)^{1/2} \end{aligned} \tag{11}$$

the maximal rms errors of prediction

$$\begin{aligned} \Delta_{\mathrm{ef}} x &= \sup_{|\psi\rangle \in \mathcal{S}} (\langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}_{\mathrm{Xf}}^2 | \psi \otimes \phi_{\mathrm{ap}} \rangle)^{1/2} \\ \Delta_{\mathrm{ef}} p &= \sup_{|\psi\rangle \in \mathcal{S}} (\langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}_{\mathrm{Pf}}^2 | \psi \otimes \phi_{\mathrm{ap}} \rangle)^{1/2} \end{aligned} \tag{12}$$

and the maximal rms disturbances

$$\begin{aligned} \Delta_{\mathrm{d}} x &= \sup_{|\psi\rangle \in \mathcal{S}} (\langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\delta}_{\mathrm{X}}^2 | \psi \otimes \phi_{\mathrm{ap}} \rangle)^{1/2} \\ \Delta_{\mathrm{d}} p &= \sup_{|\psi\rangle \in \mathcal{S}} (\langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\delta}_{\mathrm{P}}^2 | \psi \otimes \phi_{\mathrm{ap}} \rangle)^{1/2} \end{aligned} \tag{13}$$

We discussed the physical interpretation of these quantities in Appleby (1998a). The reader may confirm that this interpretation continues to be valid in the present, more general context.

It should be noted that in these definitions the supremum is only taken over all normalized initial system states. The initial apparatus state is held fixed. The quantities $\Delta_{\mathrm{ei}} x$, $\Delta_{\mathrm{ei}} p$, $\Delta_{\mathrm{ef}} x$, $\Delta_{\mathrm{ef}} p$, $\Delta_{\mathrm{d}} x$, $\Delta_{\mathrm{d}} p$ are therefore functions of the initial apparatus state.

It should also be noted that the definitions just given differ slightly from those in Appleby (1998a), in that we did not previously take the supremum over all initial system states. Some such change in the definitions is essential if the error–error and error–disturbance relationships proved in in our earlier paper for the special case of unbiased measurement processes are to be generalized so as to apply to the larger class of processes considered here. As we show in the Appendix, if one drops the requirement that the measurement be unbiased, then it is possible to find processes such that, with a suitable choice of initial state $|\psi\rangle$,

$$\langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}_{\mathrm{Xi}}^2 | \psi \otimes \phi_{\mathrm{ap}} \rangle \, \langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}_{\mathrm{Pi}}^2 | \psi \otimes \phi_{\mathrm{ap}} \rangle = 0$$

It is only when one takes the supremum over all states $|\psi\rangle$ that one gets the inequalities (21) and (22) below.

As discussed in Appleby (1998a), the quantity $(\langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}_{\mathrm{Xi}}^2 | \psi \otimes \phi_{\mathrm{ap}} \rangle)^{1/2}$ represents the rms retrodictive error in the measurement of x when the system is initially in the state $|\psi\rangle$. The quantity $\Delta_{ei}x$ (as defined above) consequently represents the maximum rms error obtained when the system is allowed to range over every possible initial state. Similarly with the quantities $\Delta_{ei}p$, $\Delta_{ef}x$, $\Delta_{ef}p$, $\Delta_{d}x$, $\Delta_{d}p$.

It is easy to think of measurement interactions for which the errors and disturbances defined in (11)–(13) are finite (with an appropriate choice of initial apparatus state). An example of such a process is the Arthurs–Kelly process (Arthurs and Kelly, 1965; Busch, 1985; Braunstein *et al.*, 1991; Stenholm, 1992; Power *et al.*, 1997; Leonhardt, 1997; Appleby, 1998a,b). In the case of the Arthurs–Kelly process, $\langle\psi \otimes \phi_{ap}|\hat{\mathbb{O}}^2|\psi \otimes \phi_{ap}\rangle$, with $\hat{\mathbb{O}}$ any error or disturbance operator, is independent of the state $|\psi\rangle$, which means that in the particular case of the Arthurs–Kelly process the definition of $\Delta_{ei}x$, $\Delta_{ei}p$, $\Delta_{ef}x$, $\Delta_{ef}p$, $\Delta_{d}x$, $\Delta_{d}p$ which are employed in this paper coincide with the definitions used in Appleby (1998a,b). It should be observed, however, that interactions for which this is true are somewhat idealized. A real measuring instrument will have a finite range. If the initial system state expectation values $\langle\psi|\hat{x}_i|\psi\rangle$ and $\langle\psi|\hat{p}_i|\psi\rangle$ are a long way outside the range of the instrument, then the errors and disturbances may not be small. Consequently, in the case of a real measuring instrument, the quantities defined in (11)–(13) may well be infinite, or at least very large. To put it another way, in the case of a real measuring instrument, these quantities do not correspond very closely to one's intuitive idea of the accuracy of and disturbance caused by a realistic measurement process. In Section 7 we show how the definitions can be modified so as to obviate this difficulty.

4. COMMUTATORS

We have, as an immediate consequence of the definitions,

$$[\hat{\varepsilon}_{Xf}, \hat{\varepsilon}_{Pf}] = i\hbar \tag{14}$$

The other commutators between the error and disturbance operators give more difficulty. This is because the retrodictive error and disturbance operators mix Heisenberg-picture observables defined at different times. It turns out, however, that it is possible to express every remaining commutator of interest in terms of commutators between one of the operators $\hat{\varepsilon}_{Xi}$, $\hat{\varepsilon}_{Pi}$, $\hat{\varepsilon}_{Xf}$, $\hat{\varepsilon}_{Pf}$, $\hat{\delta}_X$, $\hat{\delta}_P$ and one of the operators $\hat{x}_i$, $\hat{p}_i$. The significance of this result is that $\hat{x}_i$, $\hat{p}_i$ generate translations in the system phase space.

In fact

$$\begin{aligned}[\hat{\varepsilon}_{Xi}, \hat{\varepsilon}_{Pi}] &= [(\hat{\mu}_{Xf} - \hat{x}_i), (\hat{\mu}_{Pf} - \hat{p}_i)] \\ &= i\hbar - [\hat{x}_i, \hat{\mu}_{Pf}] + [\hat{p}_i, \hat{\mu}_{Xf}]\end{aligned}$$

$$= i\hbar - [\hat{x}_{\rm i}, (\hat{p}_{\rm i} + \hat{\varepsilon}_{\rm Pi})] + [\hat{p}_{\rm i}, (\hat{x}_{\rm i} + \hat{\varepsilon}_{\rm Xi})]$$

$$= -i\hbar - [\hat{x}_{\rm i}, \hat{\varepsilon}_{\rm Pi}] + [\hat{p}_{\rm i}, \hat{\varepsilon}_{\rm Xi}] \tag{15}$$

Similarly

$$[\hat{\varepsilon}_{\rm Xi}, \hat{\delta}_{\rm p}] = -i\hbar - [\hat{x}_{\rm i}, \hat{\delta}_{\rm p}] + [\hat{p}_{\rm i}, \hat{\varepsilon}_{\rm Xi}] \tag{16}$$

$$[\hat{\delta}_{\rm X}, \hat{\varepsilon}_{\rm Pi}] = -i\hbar - [\hat{x}_{\rm i}, \hat{\varepsilon}_{\rm Pi}] + [\hat{p}_{\rm i}, \hat{\delta}_{\rm X}]$$

and

$$[\hat{\varepsilon}_{\rm Xf}, \hat{\delta}_{\rm p}] = -i\hbar + [\hat{p}_{\rm i}, \hat{\varepsilon}_{\rm Xf}]$$

$$[\hat{\delta}_{\rm X}, \hat{\varepsilon}_{\rm Pf}] = -i\hbar - [\hat{x}_{\rm i}, \hat{\varepsilon}_{\rm Pf}] \tag{17}$$

5. ERROR AND ERROR–DISTURBANCE RELATIONSHIPS

We have, as an immediate consequence of (14),

$$\Delta_{\rm ef} x \, \Delta_{\rm ef} p \geq \frac{\hbar}{2} \tag{18}$$

For the remaining relationships we have to work a little harder. Let $|\psi\rangle$ be any normalized state $\in \mathscr{H}_{\rm sy}$. Let

$$\hat{D}_{xp} = \exp\left[\frac{i}{\hbar}(p\hat{x} - x\hat{p})\right]$$

be the system phase-space displacement operator, and define

$$|\psi_{xp}\rangle = \hat{D}_{xp}|\psi\rangle$$

We have

$$i\hbar \frac{\partial}{\partial x} \hat{D}_{xp} = \left(\hat{p} - \frac{1}{2} p\right)\hat{D}_{xp}, \qquad -i\hbar \frac{\partial}{\partial x} \hat{D}^{\dagger}_{xp} = \hat{D}^{\dagger}_{xp}\left(\hat{p} - \frac{1}{2} p\right)$$

$$-i\hbar \frac{\partial}{\partial p} \hat{D}_{xp} = \left(\hat{x} - \frac{1}{2} x\right)\hat{D}_{xp}, \qquad i\hbar \frac{\partial}{\partial p} \hat{D}^{\dagger}_{xp} = \hat{D}^{\dagger}_{xp}\left(\hat{x} - \frac{1}{2} x\right)$$

In view of equation (15) we then have

$$\langle \psi_{xp} \otimes \phi_{\rm ap}|[\hat{\varepsilon}_{\rm Xi}, \hat{\varepsilon}_{\rm Pi}]|\psi_{xp} \otimes \phi_{\rm ap}\rangle = -i\hbar(1 + \nabla \cdot \mathbf{v}) \tag{19}$$

where **v** is the vector

$$\mathbf{v} = \begin{pmatrix} \langle \psi_{xp} \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}_{\mathrm{Xi}} | \psi_{xp} \otimes \phi_{\mathrm{ap}} \rangle \\ \langle \psi_{xp} \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}_{\mathrm{Pi}} | \psi_{xp} \otimes \phi_{\mathrm{ap}} \rangle \end{pmatrix}$$

and ∇ is the phase space gradient operator

$$\nabla = \begin{pmatrix} \partial/\partial x \\ \partial/\partial p \end{pmatrix}$$

Now consider the box-shaped region $\mathcal{R}$ in phase space, with vertices at $(L/2, P/2)$, $(-L/2, P/2)$, $(-L/2, -P/2)$, $(L/2, -P/2)$. Let $\mathcal{C}$ be its boundary. We have

$$\begin{aligned} \Delta_{\mathrm{ei}} x \Delta_{\mathrm{ei}} p &\geq \frac{1}{2LP} \int_{\mathcal{R}} dx\, dp | \langle \psi_{xp} \otimes \phi_{\mathrm{ap}} | [\hat{\varepsilon}_{\mathrm{Xi}}, \hat{\varepsilon}_{\mathrm{Pi}}] | \psi_{xp} \otimes \phi_{\mathrm{ap}} \rangle | \\ &\geq \frac{1}{2LP} \left| \int_{\mathcal{R}} dx\, dp\, \langle \psi_{xp} \otimes \phi_{\mathrm{ap}} | [\hat{\varepsilon}_{\mathrm{Xi}}, \hat{\varepsilon}_{\mathrm{Pi}}] | \psi_{xp} \otimes \phi_{\mathrm{ap}} \rangle \right| \\ &\geq \frac{\hbar}{2} \left(1 - \frac{1}{LP} \left| \int_{\mathcal{R}} dx\, dp\, \nabla \cdot \mathbf{v} \right| \right) \\ &= \frac{\hbar}{2} \left(1 - \frac{1}{LP} \left| \int_{\mathcal{C}} ds\, \mathbf{n} \cdot \mathbf{v} \right| \right) \\ &\geq \frac{\hbar}{2} \left(1 - \frac{2}{L} \Delta_{\mathrm{ei}} x - \frac{2}{P} \Delta_{\mathrm{ei}} p \right) \end{aligned} \tag{20}$$

where ds is the line element and $\mathbf{n}$ is the outward-pointing unit normal along $\mathcal{C}$. Taking the limit as $L, P \to \infty$, we deduce

$$\Delta_{\mathrm{ei}} x\, \Delta_{\mathrm{ei}} p \geq \frac{\hbar}{2} \tag{21}$$

whenever the left-hand side is defined (i.e., whenever it is not of the form $0 \times \infty$).

Starting from (16) and (17) we deduce, by essentially the same argument,

$$\begin{aligned} \Delta_{\mathrm{ei}} x \Delta_{\mathrm{d}} p &\geq \frac{\hbar}{2}, \qquad \Delta_{\mathrm{ef}} x\, \Delta_{\mathrm{d}} p \geq \frac{\hbar}{2} \\ \Delta_{\mathrm{ei}} p \Delta_{\mathrm{d}} x &\geq \frac{\hbar}{2}, \qquad \Delta_{\mathrm{ef}} p\, \Delta_{\mathrm{d}} x \geq \frac{\hbar}{2} \end{aligned} \tag{22}$$

whenever the products are defined.

It should be noted that although the relationships proved in this section have the same form as the corresponding relationships proved in Appleby (1998a), they do not have the same content, since the quantities $\Delta_{\mathrm{ei}} x$, $\Delta_{\mathrm{ei}} p$, $\Delta_{\mathrm{ef}} x$, $\Delta_{\mathrm{ef}} p$, $\Delta_{\mathrm{d}} x$, $\Delta_{\mathrm{d}} p$ appearing in them are not defined in the same way (in our previous publications we did not take a supremum over all normalized initial system states when defining the errors and disturbances; see Section 3 above, and Section 6 immediately following).

In the Introduction we remarked that in the case of the retrodictively unbiased measurement processes considered in Appleby (1998a), the Arthurs–Kelly relationship [relation (6) above] is a consequence of the retrodictive error relationship. It is an interesting question, which we have not as yet been able to resolve, whether it is possible to deduce an Arthurs–Kelly type bound from (21), applying to the much more general class of measurement processes considered in this paper.

6. UNBIASED MEASUREMENTS

Suppose that the measurement process is retrodictively unbiased, in the sense that

$$\langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}_{\mathrm{Xi}} | \psi \otimes \phi_{\mathrm{ap}} \rangle = \langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}_{\mathrm{Pi}} | \psi \otimes \phi_{\mathrm{ap}} \rangle = 0$$

uniformly, for all $|\psi\rangle \in \mathcal{H}_{\mathrm{sy}}$ (but fixed $|\phi_{\mathrm{ap}}\rangle$). Then the vector $\mathbf{v}$ appearing on the right-hand side of (19) is identically zero, and we have

$$\langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}^2_{\mathrm{Xi}} | \psi \otimes \phi_{\mathrm{ap}} \rangle \langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}^2_{\mathrm{Pi}} | \psi \otimes \phi_{\mathrm{ap}} \rangle \geq \frac{\hbar^2}{4}$$

uniformly, for all $|\psi\rangle \in \mathcal{H}_{\mathrm{sy}}$.

Suppose, in addition, that the measurement is predictively unbiased:

$$\langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}_{\mathrm{Xf}} | \psi \otimes \phi_{\mathrm{ap}} \rangle = \langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}_{\mathrm{Pf}} | \psi \otimes \phi_{\mathrm{ap}} \rangle = 0$$

for all $|\psi\rangle$. Then we have, by a similar argument,

$$\langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}^2_{\mathrm{Xi}} | \psi \otimes \phi_{\mathrm{ap}} \rangle \, \langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\delta}^2_{\mathrm{P}} | \psi \otimes \phi_{\mathrm{ap}} \rangle \geq \hbar^2/4$$

$$\langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}^2_{\mathrm{Xf}} | \psi \otimes \phi_{\mathrm{ap}} \rangle \, \langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\delta}^2_{\mathrm{P}} | \psi \otimes \phi_{\mathrm{ap}} \rangle \geq \hbar^2/4$$

$$\langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}^2_{\mathrm{Pi}} | \psi \otimes \phi_{\mathrm{ap}} \rangle \, \langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\delta}^2_{\mathrm{X}} | \psi \otimes \phi_{\mathrm{ap}} \rangle \geq \hbar^2/4$$

$$\langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\varepsilon}^2_{\mathrm{Pf}} | \psi \otimes \phi_{\mathrm{ap}} \rangle \, \langle \psi \otimes \phi_{\mathrm{ap}} | \hat{\delta}^2_{\mathrm{X}} | \psi \otimes \phi_{\mathrm{ap}} \rangle \geq \hbar^2/4$$

uniformly, for all $|\psi\rangle$.

These are the results which we proved in Appleby (1998a) by a different method.

7. MEASUREMENTS WITH A FINITE RANGE

Real measuring instruments are only designed to be used for a limited set of initial system states. For such an instrument one expects the maximal rms errors and disturbances defined in (11)–(13) to be infinite, or at least very large. This is because the supremum is taken over every possible initial system state, including those states for which the expected values of $\hat{x}$ and $\hat{p}$ are far outside the range of the instrument. It follows that the quantities defined in (11)–(13) are poor indicators of the accuracies and disturbances to be expected when the instrument is used in the manner in which it was designed to be used. In the case of a real measuring instrument, what interests us are the maximum errors and disturbances obtained for a *limited class* of initial system states—namely, the class on which the instrument was designed to make measurements. In this section we discuss an alternative definition of the errors and disturbances which is more appropriate to such a case.

Suppose that the instrument is designed to be accurate for initial system states $|\psi\rangle$ such that

$$x_0 - \tfrac{1}{2}L \leq \langle\psi|\hat{x}|\psi\rangle \leq x_0 + \tfrac{1}{2}L, \qquad p_0 - \tfrac{1}{2}P \leq \langle\psi|\hat{p}|\psi\rangle \leq p_0 + \tfrac{1}{2}P$$

and

$$\Delta x \leq \sigma, \qquad \Delta p \leq \tau$$

for fixed constants x_0, p_0, L, P, σ, τ such that $\sigma\tau \geq \hbar/2$. Let $\mathcal{S}'$ be the set of normalized states $\varepsilon\ \mathcal{H}_{\text{sy}}$ which satisfy these conditions. The errors and disturbances appropriate for the description of this instrument are obtained by taking the supremum over all normalized states $|\psi\rangle\ \varepsilon\ \mathcal{S}'$:

$$\begin{aligned} \Delta'_{\text{ei}}x &= \sup_{|\psi\rangle\in\mathcal{S}'} (\langle\psi\otimes\phi_{\text{ap}}|\hat{\varepsilon}^2_{\text{Xi}}|\psi\otimes\phi_{\text{ap}}\rangle)^{1/2} \\ \Delta'_{\text{ei}}p &= \sup_{|\psi\rangle\in\mathcal{S}'} (\langle\psi\otimes\phi_{\text{ap}}|\hat{\varepsilon}^2_{\text{Pi}}|\psi\otimes\phi_{\text{ap}}\rangle)^{1/2} \end{aligned} \tag{23}$$

$$\begin{aligned} \Delta'_{\text{ef}}x &= \sup_{|\psi\rangle\in\mathcal{S}'} (\langle\psi\otimes\phi_{\text{ap}}|\hat{\varepsilon}^2_{\text{Xf}}|\psi\otimes\phi_{\text{ap}}\rangle)^{1/2} \\ \Delta'_{\text{ef}}p &= \sup_{|\psi\rangle\in\mathcal{S}'} (\langle\psi\otimes\phi_{\text{ap}}|\hat{\varepsilon}^2_{\text{Pf}}|\psi\otimes\phi_{\text{ap}}\rangle)^{1/2} \end{aligned} \tag{24}$$

$$\begin{aligned} \Delta'_{\text{d}}x &= \sup_{|\psi\rangle\in\mathcal{S}'} (\langle\psi\otimes\phi_{\text{ap}}|\hat{\delta}^2_{\text{X}}|\psi\otimes\phi_{\text{ap}}\rangle)^{1/2} \\ \Delta'_{\text{d}}p &= \sup_{|\psi\rangle\in\mathcal{S}'} (\langle\psi\otimes\phi_{\text{ap}}|\hat{\delta}^2_{\text{P}}|\psi\otimes\phi_{\text{ap}}\rangle)^{1/2} \end{aligned} \tag{25}$$

It follows from (14) that

$$\Delta'_{\text{ef}} x \Delta'_{\text{ef}} p \geq \frac{\hbar}{2}$$

Turning to the retrodictive error relationship, let $|\psi\rangle$ be any normalized state $\in \mathcal{H}_{\text{sy}}$ such that

$$\langle\psi|\hat{x}|\psi\rangle = x_0, \qquad \langle\psi|\hat{p}|\psi\rangle = p_0$$

and

$$\Delta x \leq \sigma, \qquad \Delta p \leq \tau$$

Let $\mathcal{R}$ be the box-shaped region of phase space with vertices $(x_0 + L/2, p_0 + P/2)$, $(x_0 - L/2, p_0 + P/2)$, $(x_0 - L/2, p_0 - P/2)$, $(x_0 + L/2, p_0 - P/2)$. Then $|\psi_{xp}\rangle \in \mathcal{S}'$ for all $(x,p) \in \mathcal{R}$. We can now use an argument analogous to the one leading to (20) to deduce

$$\Delta'_{\text{ei}} x \Delta'_{\text{ei}} p \geq \frac{\hbar}{2} \left(1 - \frac{2}{L} \Delta'_{\text{ei}} x - \frac{2}{P} \Delta'_{\text{ei}} P \right)$$

which can alternatively be written

$$\left(\Delta'_{\text{ei}} x + \frac{\hbar}{P} \right) \left(\Delta'_{\text{ei}} p + \frac{\hbar}{L} \right) \geq \frac{\hbar}{2} \left(1 + \frac{2\hbar}{LP} \right) \tag{26}$$

If $P\,\Delta'_{\text{ei}} x$, $L\,\Delta'_{\text{ei}} p$, and LP are all $\gg \hbar$, we have the approximate relation

$$\Delta'_{\text{ei}} x \Delta'_{\text{ei}} p \gtrsim \frac{\hbar}{2}$$

One expects this approximate form of the retrodictive error relationship to be valid in most situations of practical interest. However, it is not always valid (see the Appendix for a counterexample).

Starting from (16) and (17), we can derive in a similar manner

$$\begin{aligned} \left(\Delta'_{\text{ei}} x + \frac{\hbar}{P} \right) \left(\Delta'_{\text{d}} p + \frac{\hbar}{L} \right) &\geq \frac{\hbar}{2} \left(1 + \frac{2\hbar}{LP} \right) \\ \left(\Delta'_{\text{ei}} p + \frac{\hbar}{L} \right) \left(\Delta'_{\text{d}} x + \frac{\hbar}{P} \right) &\geq \frac{\hbar}{2} \left(1 + \frac{2\hbar}{LP} \right) \end{aligned} \tag{27}$$

and

$$\Delta'_{\text{ef}} x \left(\Delta'_{\text{d}} p + \frac{\hbar}{L} \right) \geq \frac{\hbar}{2}$$
$$\Delta'_{\text{ef}} p \left(\Delta'_{\text{d}} x + \frac{\hbar}{P} \right) \geq \frac{\hbar}{2} \tag{28}$$

8. CONCLUDING REMARKS

The commonest method of describing the spread of a statistical distribution, in terms of the variance—the method employed in this paper, in other words—is subject to certain limitations. In recent years there has accordingly been some interest in devising alternative approaches. One approach is that involving parameter-based uncertainty relationships (Hilgevoord and Uffink, 1990; Braunstein *et al.*, 1996). Another approach is that involving entropic uncertainty relationships (Busch and Lahti, 1989; Martens and de Muynck, 1990a,b, 1992; Ban, 1997; Buzek *et al.*, 1995a,b). It would be interesting to see if either of these approaches can be used to develop the results obtained in this paper.

We should also remark that in this paper we have made no use of the mathematical theory based on the concept of a POVM and an unsharp observable (Prugovecki, 1973, 1976a,b, 1978, 1984; Davies, 1976; Holevo, 1982; Busch, 1985; Busch and Lahti, 1989; Martens and de Muynck, 1990a,b, 1992; de Muynck *et al.*, 1994; Busch *et al.*, 1995; Ban, 1997). There were certain advantages in proceeding in this way. One advantage was that it enabled us to circumvent the difficulties which have been identified by Uffink (1994), as we discussed in Appleby (1998a). Also, we share the view of Englert and Wódkiewicz (1995) that the underlying intrinsic observables should be regarded as "the heart of the matter." One of the advantages of the approach adopted here is that it places the emphasis on these intrinsic observables, as opposed to (in the words of Englert and Wódkiewicz) a "mathematical representation of the statistical information gathered." Nevertheless, the theory of POVM's is clearly an important and very powerful way of analyzing simultaneous measurement processes. We certainly do not mean to set up the approach taken in this paper as an *alternative* to the approach based on POVMs. We merely wish to stress the point made by Englert and Wódkiewicz, that POVMs and unsharp observables should be regarded as secondary mathematical constructs, rather than as fundamental physical concepts which need to be posited from the outset. We hope to return to this question in a future publication, in which we will show how the concept of an unsharp observable naturally emerges from the approach taken in this paper.

APPENDIX

The purpose of this appendix is to explain why we defined the errors and disturbances by taking the supremum over every normalized initial system state, as in (11)–(13), or a subset of them, as in (23)–(25). The reason is that there exist processes such that (for example)

$$\langle \psi \otimes \phi_{ap} | \hat{\varepsilon}^2_{Xi} | \psi \otimes \phi_{ap} \rangle \langle \psi \otimes \phi_{ap} | \hat{\varepsilon}^2_{Pi} | \psi \otimes \phi_{ap} \rangle = 0$$

for certain choices of initial system state $|\psi\rangle$ and initial apparatus state $|\phi_{ap}\rangle$. It is only when one takes the appropriate supremum that one gets the inequalities (21) and (22) or (26)–(28).

Consider, for example, the measurement interaction described by the evolution operator

$$\hat{U} = \exp\left[-\frac{i\pi}{2\hbar} (\hat{x}\hat{\pi}_X - \hat{\mu}_X \hat{p}) \right]$$

where $\hat{\pi}_X$ is a momentum canonically conjugate to the pointer observable $\hat{\mu}_X$. $\hat{U}$ is a rotation operator in $xp\mu_X\pi_X$ space. It takes $\hat{\mu}_X$ onto $\hat{x}$ and $\hat{x}$ onto $-\hat{\mu}_X$:

$$\begin{pmatrix} \hat{x}_f \\ \hat{\mu}_{Xf} \end{pmatrix} = \hat{U}^\dagger \begin{pmatrix} \hat{x} \\ \hat{\mu}_X \end{pmatrix} \hat{U} = \begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{\mu}_X \end{pmatrix} = \begin{pmatrix} -\hat{\mu}_X \\ \hat{x} \end{pmatrix}$$

Similarly

$$\begin{pmatrix} \hat{p}_f \\ \hat{\pi}_{Xf} \end{pmatrix} = \begin{pmatrix} -\hat{\pi}_X \\ \hat{p} \end{pmatrix}$$

$\hat{\mu}_P$ is unaffected by the interaction. Referring back to the definitions (8)–(10), we deduce

$$\hat{\varepsilon}_{Xi} = 0, \qquad \hat{\varepsilon}_{Xf} = \hat{\mu}_X + \hat{x}, \qquad \hat{\delta}_X = -\hat{\mu}_X - \hat{x}$$

$$\hat{\varepsilon}_{Pi} = \hat{\mu}_P - \hat{p}, \qquad \hat{\varepsilon}_{Pf} = \hat{\mu}_P + \hat{\pi}_X, \qquad \hat{\delta}_P = -\hat{\pi}_X - \hat{p}$$

Since $\hat{\mu}_{Xf} = \hat{x}_i$ the process effects a perfectly accurate retrodiction of position, and this is reflected in the fact that $\Delta_{ei}x = 0$. On the other hand, the momentum pointer is unaffected by the interaction: $\hat{\mu}_{Pf} = \hat{\mu}_{Pi}$. This means that the process is not really measuring the momentum at all. We accordingly find $\Delta_{ei}\,p = \infty$. If we use the alternative definition of (23), then we find

$$\Delta'_{ei} p \geq \frac{P}{2}$$

which is again consistent with the fact, that so far as momentum is concerned, the process hardly counts as a measurement. Nevertheless, from the fact that

$$\langle \psi \otimes \phi_{ap} | \hat{\varepsilon}^2_{Pi} | \psi \otimes \phi_{ap} \rangle = (\Delta \mu_P)^2 + (\Delta p)^2 + (\langle \phi_{ap} | \hat{\mu}_P | \phi_{ap} \rangle - \langle \psi | \hat{p} | \psi \rangle)^2$$

we see, that by appropriately choosing $|\psi\rangle$ and $|\phi_{ap}\rangle$, $\langle \hat{\varepsilon}^2_{Pi} \rangle$ can be made arbitrarily small. Moreover, the product $\langle \hat{\varepsilon}^2_{Xi} \rangle \langle \hat{\varepsilon}^2_{Pi} \rangle$ will be zero whenever $\langle \hat{\varepsilon}^2_{Pi} \rangle$ is finite.

It is not surprising that $\langle \hat{\varepsilon}^2_{Pi} \rangle$ is small for certain choices of initial state. Consider, for example, the classical situation, where one has a classical ammeter whose needle is stuck at the 1-A position. Then the meter will, of course, give exactly the right reading if one uses it to measure a 1-A current.

REFERENCES

Appleby, D. M. (1998a). *International Journal of Theoretical Physics*, **37**, 1491.
Appleby, D. M. (1998b). *Journal of Physics A*, **31**, 6419.
Arthurs, E., and Goodman, M. S. (1988). *Physical Review Letters*, **60**, 2447.
Arthurs, E., and Kelly, J. L. (1965). *Bell System Technical Journal*, **44**, 725.
Ballentine, L. F. (1970). *Reviews of Modern Physics*, **42**, 358.
Ban, M. (1997). *International Journal of Theoretical Physics*, **36**, 2583.
Bohm, D. (1951). *Quantum Theory*, Prentice-Hall, New York.
Braginsky, V. B., and Khalili, F. Ya. (1992). *Quantum Measurement*, K. S. Thorne, ed., Cambridge University Press, Cambridge.
Braunstein, S. L., Caves, C. M., and Milburn, G. J. (1991). *Physical Review A*, **43**, 1153.
Braunstein, S. L., Caves, C. M., and Milburn, G. J. (1996). *Annals of Physics*, **247**, 135.
Busch, P. (1985). *International Journal of Theoretical Physics*, **24**, 63.
Busch, P., and Lahti, P. J. (1989). *Foundations of Physics*, **19**, 633.
Busch, P., Grabowski, M., and Lahti, P. J. (1995). *Operational Quantum Physics*, Springer-Verlag, Berlin.
Bužek, V., Keitel, C. H., and Knight, P. L. (1995a). *Physical Review A*, **51**, 2575.
Bužek, V., Keitel, C. H., and Knight, P. L. (1995b). *Physical Review A*, **51**, 2594.
Davies, E. B. (1976). *Quantum Theory of Open Systems*, Academic Press, New York.
de Muynck, W. M., de Baere, W., and Martens, H. (1994). *Foundations of Physics*, **24**, 1589.
Englert, B.-G., and Wódkiewicz, K. (1995). *Physical Review A*, **51**, R2661.
Heisenberg, W. (1927). *Zeitschrift für Physik*, **43**, 172 [Reprinted in *Quantum Theory and Measurement*, J. A. Wheeler and W. H. Zurek, eds., Princeton University Press, Princeton, New Jersey, 1983].
Heisenberg, W. (1930). *The Physical Principles of the Quantum Theory*, University of Chicago Press, Chicago.
Hilgevoord, J., and Uffink, J. (1990). In *Sixty Two Years of Uncertainty*, A. I. Miller, ed., Plenum Press, New York.
Holevo, A. S. (1982). *Probabilistic and Statistical Aspects of Quantum Theory*, North-Holland, Amsterdam.

Leonhardt, U. (1997). *Measuring the Quantum State of Light*, Cambridge University Press, Cambridge.
Leonhardt, U., and Paul, H. (1993a). *Journal of Modern Optics*, **40**, 1745.
Leonhardt, U., and Paul, H. (1993b). *Physical Review A*, **48**, 4598.
Leonhardt, U., and Paul, H. (1995a). *Optics Communications*, **119**, 296.
Leonhardt, U., and Paul, H. (1995b). *Progress in Quantum Electronics*, **19**, 89.
Margenau, H. (1958). *Philosophy of Science*, **25**, 23.
Martens, H., and de Muynck, W. M. (1990a). *Foundations of Physics*, **20**, 255.
Martens, H., and de Muynck, W. M. (1990b). *Foundations of Physics*, **20**, 357.
Martens, H., and de Muynck, W. M. (1992). *Journal of Physics A*, **25**, 4887.
Power, W. L., Tan, S. M., and Wilkens, M. (1997). *Journal of Modern Optics*, **44**, 2591.
Prugovečki, E. (1973). *Foundations of Physics*, **3**, 3.
Prugovečki, E. (1976a). *Journal of Mathematical Physics*, **17**, 517.
Prugovečki, E. (1976b). *Journal of Mathematical Physics*, **17**, 1673.
Prugovečki, E. (1978). *Annals of Physics*, **110**, 102.
Prugovečki, E. (1984). *Stochastic Quantum Mechanics and Quantum Spacetime*, Reidel, Dordrecht.
Raymer, M. G. (1994). *American Journal of Physics*, **62**, 986.
Stenholm, S. (1992). *Annals of Physics*, **218**, 233.
Törma, P., Stenholm, S., and Jex, I. (1995). *Physical Review A*, **52**, 4812.
Uffink, J. (1994). *International Journal of Theoretical Physics*, **33**, 199.
Wódkiewicz, K. (1987). *Physics Letters A*, **124**, 207.