

Optimal Joint Measurements of Position and Momentum

D. M. Appleby¹

Received July 2, 1998

The distribution of measured values for maximally accurate, unbiased simultaneous measurements of position and momentum is investigated. It is shown, that if the measurement is retrodictively optimal, then the distribution of results is given by the initial state Husimi function (or Q -representation). If the measurement is predictively optimal, then the distribution of results is related to the final state anti-Husimi function (or P -representation). The significance of this universal property for the interpretation of the Husimi function is discussed.

1. INTRODUCTION

There is currently some interest in simultaneous measurements of position and momentum (Arthurs and Kelly, 1965; Wódkiewicz, 1984, 1986, 1987; Braunstein *et al.*, 1991; Stenholm, 1992; Halliwell, 1992; Burak and Wódkiewicz, 1992; Leonhardt and Paul, 1993, 1995; Törma *et al.*, 1995; Englert and Wódkiewicz, 1995; Busch *et al.*, 1995; Power *et al.*, 1997; Banaszek and Wódkiewicz, 1997; Ban, 1997; Leonhardt, 1997). Measurements of this kind have an immediate, technical relevance to the field of quantum optics. They also have a rather more general, conceptual relevance to the problem of understanding the classical limit.

In Appleby (1998a–c) we discussed the accuracy of such measurements. We began with Braginsky and Khalili's (1992) analysis of single measurements of x only, and extended it to the case of simultaneous measurements of x and p together. We identified two types of error: the retrodictive (or determinative) errors $\Delta_{ei}x$, $\Delta_{ei}p$, and the predictive (or preparative) errors $\Delta_{ef}x$, $\Delta_{ef}p$. We showed that, subject to some rather unrestrictive assumptions

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

regarding the nature of the measurement process, they satisfy the retrodictive error relationship

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

and the predictive error relationship

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

In the following we address the question: What (if anything) can be said about the distribution of measured values in those cases where the lower bound set by one of these inequalities is actually achieved?

We begin, in Section 2, by considering measurements which are retrodictively optimal. We define a retrodictively optimal measurement to be any measurement belonging to the class of processes defined in Appleby (1998c) which minimizes the product of retrodictive errors (so that $\Delta_{ei}x\Delta_{ei}p = \hbar/2$), and which is retrodictively unbiased [so that the systematic errors of retrodiction are zero—see Eq. (1) below]. We show that for such measurements the distribution of measured values is always given by the initial-system-state Husimi function, or Q -function (Husimi, 1940; Hillery *et al.*, 1984; Lee, 1995; Leonhardt, 1997).

A number of related results have been obtained by other authors. In the case of the Arthurs–Kelly process several authors (Arthurs and Kelly, 1965; Braunstein *et al.*, 1991; Stenholm, 1992; Leonhardt, 1997) have shown that the Husimi function describes the distribution of measured values for certain choices of initial apparatus state. Leonhardt and Paul (1993, 1995) have shown that the same is true for a number of other processes. However, none of these authors relate the distribution of measured values to the accuracy of the measurement process.

Busch (1985) and Appleby (1998b) have analyzed the Arthurs–Kelly process in terms of the accuracy of the measured values, and have shown that the Husimi function gives the distribution of results in precisely those cases where the measurement is retrodictively “optimal” or “best.”

Wódkiewicz has proposed an “operational” approach to the theory of phase space distributions (Wódkiewicz, 1984, 1986, 1987; Burak and Wódkiewicz, 1992; Englert and Wódkiewicz, 1995; Banaszek and Wódkiewicz, 1997; Ban, 1997; Marchioli *et al.*, 1997). If one takes the filter reference state (or “quantum ruler”) used to define the operational distribution to be a squeezed vacuum state, and a minimum uncertainty state for \hat{x} and \hat{p} , then one obtains the Husimi function. It could be said that the Husimi function is the operational distribution corresponding to the case when the quantum ruler is most exactly

and finely calibrated—a fact which obviously ties in with the result which we will prove in Section 2.

However, the result which is most similar to ours is the one obtained by Ali and Prugovečki (1977), working within the framework of the approach based on POVMs (positive operator-valued measures) and unsharp observables (Prugovečki, 1976a, b, 1978, 1984; Holevo, 1982; Busch, 1985; Busch *et al.*, 1995; Ban, 1997). Ali and Prugovečki prove a general theorem. However, it is considerably less general than the theorem which we will prove in Section 2. In the first place, the definition of the experimental accuracy which is employed by Ali and Prugovečki is only valid for a limited class of measurement processes, as is shown in the appendix to Appleby (1998b). The validity of their result is further restricted by the fact that they only prove it for processes which are Galilean covariant. In Section 2 we prove the proposition under assumptions which are much less restrictive (in fact, it turns out that a retrodictively optimal measurement process always is Galilean covariant; however, that is a consequence of our argument—we do not begin by assuming it). It may also be worth remarking that our way of analyzing simultaneous measurement processes is rather different from the approach of Ali and Prugovečki. In particular, the objections raised by Uffink (1994) do not apply to our arguments.

In Section 3 we go on to consider predictively optimal measurements—*i.e.*, measurements of the type defined in Appleby (1998c) which minimize the product of predictive errors (so that $\Delta_{efx} \Delta_{efp} = \hbar/2$). We show that in the case of such a measurement the distribution of results is related to the final-state anti-Husimi function, or *P*-function (Glauber, 1963; Sudarshan, 1963; Hillery *et al.*, 1984; Lee, 1995; Leonhardt, 1997). This result extends a result proved in Appleby (1998b) for the particular case of the Arthurs–Kelly process to the general class of measurement processes defined in Appleby (1998c).

In Section 4 we conclude by discussing the bearing of our results on the interpretation of the Husimi function. In Section 2 we show that the Husimi function describes the outcome of *any* retrodictively optimal process. In other words, the Husimi function has a universal significance. We will argue that this lends some support to the idea that the Husimi function is the quantum mechanical entity which most nearly resembles the classical concept of the “real” or “objective” distribution describing an ensemble of identically prepared systems.

2. RETRODICTIVELY OPTIMAL MEASUREMENTS

We will say that a simultaneous measurement process of the kind defined in Appleby (1998c) is *retrodictively optimal* if the following hold:

1. The process is retrodictively unbiased, so that

$$\langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{X_i} | \psi \otimes \phi_{\text{ap}} \rangle = \langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{P_i} | \psi \otimes \phi_{\text{ap}} \rangle = 0 \quad (1)$$

for all $|\psi\rangle \in \mathcal{H}_{\text{sy}}$.

2. The product of retrodictive errors achieves its lower bound, so that

$$\Delta_{\epsilon_i X} \Delta_{\epsilon_i P} = \frac{\hbar}{2} \quad (2)$$

Here and in the sequel we employ the notation and terminology of Appleby (1998c). Thus, $|\psi\rangle \in \mathcal{H}_{\text{sy}}$ and $|\phi_{\text{ap}}\rangle \in \mathcal{H}_{\text{ap}}$ are the initial states of the system and apparatus, respectively. $\hat{\epsilon}_{X_i}$, $\hat{\epsilon}_{P_i}$ are the retrodictive error operators. $\Delta_{\epsilon_i X}$, $\Delta_{\epsilon_i P}$ are the maximal rms errors of retrodiction.

In Appleby (1998b) we considered the special case of the Arthurs–Kelly process. In that case one has the commutation relation

$$[\hat{\epsilon}_{X_i}, \hat{\epsilon}_{P_i}] = -i\hbar \quad (3)$$

This relationship, and the condition of Eq. (2), together imply Eq. (1). In the general case, however, it is necessary to impose as a separate condition the requirement that the measurement be retrodictively unbiased.

In the general case the commutation relationship of Eq. (3) cannot be assumed. However, it was shown in Appleby (1998c) that Eq. (1) implies the weaker statement

$$\langle \psi \otimes \phi_{\text{ap}} | [\hat{\epsilon}_{X_i}, \hat{\epsilon}_{P_i}] | \psi \otimes \phi_{\text{ap}} \rangle = -i\hbar \quad (4)$$

for every normalized $|\psi\rangle \in \mathcal{H}_{\text{sy}}$ (but fixed $|\phi_{\text{ap}}\rangle$). It turns out that this is enough to prove that the distribution of measured values is given by the initial system state Husimi function for any retrodictively optimal process. However, the fact that we can no longer assume the commutation relationship of Eq. (3) means that the proof of this statement is less straightforward than the proof given in Appleby (1998b) for the special case of the Arthurs–Kelly process.

In view of Eqs. (2) and (4) we have

$$\langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{X_i}^2 | \psi \otimes \phi_{\text{ap}} \rangle \langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{P_i}^2 | \psi \otimes \phi_{\text{ap}} \rangle = \frac{\hbar^2}{4} \quad (5)$$

for every normalized $|\psi\rangle \in \mathcal{H}_{\text{sy}}$. We deduce:

Lemma 1. Given any retrodictively optimal measurement process with initial apparatus state $|\phi_{\text{ap}}\rangle$, there exists a fixed number λ_i such that

$$\langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{\text{Xi}}^2 | \psi \otimes \phi_{\text{ap}} \rangle = \frac{\lambda_i^2}{2}$$

$$\langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{\text{Pi}}^2 | \psi \otimes \phi_{\text{ap}} \rangle = \frac{\hbar^2}{2\lambda_i^2}$$

for every normalized $|\psi\rangle \in \mathcal{H}_{\text{sy}}$.

Remark. We will refer to λ_i as the retrodictive spatial resolution of the measurement.

Proof. For each normalized $|\psi\rangle \in \mathcal{H}_{\text{sy}}$ define the number λ_ψ by

$$\lambda_\psi = (2\langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{\text{Xi}}^2 | \psi \otimes \phi_{\text{ap}} \rangle)^{1/2}$$

In view of Eq. (5) we then have

$$(\langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{\text{Pi}}^2 | \psi \otimes \phi_{\text{ap}} \rangle)^{1/2} = \frac{\hbar}{\sqrt{2}\lambda_\psi}$$

We have from the definitions of $\Delta_{\text{ei}x}$, $\Delta_{\text{ei}p}$ (Appleyby, 1998c)

$$\Delta_{\text{ei}x} = \sup_{|\psi\rangle \in \mathcal{S}} (\langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{\text{Xi}}^2 | \psi \otimes \phi_{\text{ap}} \rangle)^{1/2} = \frac{\sup_{|\psi\rangle \in \mathcal{S}} (\lambda_\psi)}{\sqrt{2}}$$

and

$$\Delta_{\text{ei}p} = \sup_{|\psi\rangle \in \mathcal{S}} (\langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{\text{Pi}}^2 | \psi \otimes \phi_{\text{ap}} \rangle)^{1/2} = \frac{\hbar}{\sqrt{2} \inf_{|\psi\rangle \in \mathcal{S}} (\lambda_\psi)}$$

where \mathcal{S} denotes the unit sphere in the system state space. In view of Eq. (2) it then follows that

$$\inf_{|\psi\rangle \in \mathcal{S}} (\lambda_\psi) = \sup_{|\psi\rangle \in \mathcal{S}} (\lambda_\psi)$$

which means that λ_ψ must be constant. ■

We next define the operators

$$\begin{aligned} \hat{c}_{\lambda_i} &= \frac{1}{\sqrt{2}} \left(\frac{1}{\lambda_i} \hat{\epsilon}_{\text{Xi}} - \frac{i\lambda_i}{\hbar} \hat{\epsilon}_{\text{Pi}} \right) \\ \hat{c}_{\lambda_i}^\dagger &= \frac{1}{\sqrt{2}} \left(\frac{1}{\lambda_i} \hat{\epsilon}_{\text{Xi}} + \frac{i\lambda_i}{\hbar} \hat{\epsilon}_{\text{Pi}} \right) \end{aligned} \quad (6)$$

In the general case we cannot assume the commutation relation of Eq. (3). It follows that \hat{c}_{λ_i} , $\hat{c}_{\lambda_i}^\dagger$ are not, in general, ladder operators. We do, however, have the relationship of Eq. (4), and this is enough to prove the following.

Lemma 2. Given any retrodictively optimal measurement process with initial apparatus state $|\phi_{\text{ap}}\rangle$ and retrodictive spatial resolution λ_i , let \hat{c}_{λ_i} be the operator defined by Eq. (6). Then

$$\hat{c}_{\lambda_i}|\psi\rangle \otimes |\phi_{\text{ap}}\rangle = 0$$

for every $|\psi\rangle \in \mathcal{H}_{\text{sy}}$.

Proof. Given any normalized system state $|\psi\rangle$, let $\alpha, \beta \in \mathbb{R}$ be the real and imaginary parts of $\langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{X_i} \hat{\epsilon}_{P_i} | \psi \otimes \phi_{\text{ap}} \rangle$:

$$\langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{X_i} \hat{\epsilon}_{P_i} | \psi \otimes \phi_{\text{ap}} \rangle = \alpha + i\beta \quad (7)$$

We have

$$\begin{aligned} (\alpha^2 + \beta^2)^{1/2} &= |\langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{X_i} \hat{\epsilon}_{P_i} | \psi \otimes \phi_{\text{ap}} \rangle| \\ &= \|\hat{\epsilon}_{X_i}|\psi\rangle \otimes |\phi_{\text{ap}}\rangle\| \cdot \|\hat{\epsilon}_{P_i}|\psi\rangle \otimes |\phi_{\text{ap}}\rangle\| = \frac{\hbar}{2} \end{aligned}$$

where

$$\begin{aligned} \|\hat{\epsilon}_{X_i}|\psi\rangle \otimes |\phi_{\text{ap}}\rangle\| &= \left(\langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{X_i}^2 | \psi \otimes \phi_{\text{ap}} \rangle \right)^{1/2} = \frac{\lambda_i}{\sqrt{2}} \\ \|\hat{\epsilon}_{P_i}|\psi\rangle \otimes |\phi_{\text{ap}}\rangle\| &= \left(\langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{P_i}^2 | \psi \otimes \phi_{\text{ap}} \rangle \right)^{1/2} = \frac{\hbar}{\sqrt{2}\lambda_i} \end{aligned}$$

are the norms of the vectors $\hat{\epsilon}_{X_i}|\psi\rangle \otimes |\phi_{\text{ap}}\rangle$, $\hat{\epsilon}_{P_i}|\psi\rangle \otimes |\phi_{\text{ap}}\rangle$.

In view of Eq. (4) we also have

$$-i\hbar = \langle \psi \otimes \phi_{\text{ap}} | [\hat{\epsilon}_{X_i}, \hat{\epsilon}_{P_i}] | \psi \otimes \phi_{\text{ap}} \rangle = 2i\beta$$

Consequently, $\alpha = 0$ and $\beta = -\hbar/2$. We then have

$$|\langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{X_i} \hat{\epsilon}_{P_i} | \psi \otimes \phi_{\text{ap}} \rangle| = \frac{\hbar}{2} = \|\hat{\epsilon}_{X_i}|\psi\rangle \otimes |\phi_{\text{ap}}\rangle\| \cdot \|\hat{\epsilon}_{P_i}|\psi\rangle \otimes |\phi_{\text{ap}}\rangle\|$$

Now it is generally true in any Hilbert space, that two vectors $|\Psi_1\rangle$, $|\Psi_2\rangle$ having the property

$$|\langle \Psi_1 | \Psi_2 \rangle| = \|\Psi_1\| \cdot \|\Psi_2\|$$

must be parallel. Hence

$$\hat{\epsilon}_{P_i}|\psi\rangle \otimes |\phi_{\text{ap}}\rangle = \gamma \hat{\epsilon}_{X_i}|\psi\rangle \otimes |\phi_{\text{ap}}\rangle$$

for some $\gamma \in \mathbb{C}$. Inserting this result into Eq. (7), we find

$$\gamma = -\frac{i\hbar}{\lambda_i^2}$$

The claim follows. ■

Now let

$$\rho(\mu_{Xf}, \mu_{Pf}) = \int dx_f dy_{f1} \dots dy_{fn} |\langle x_f, \mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn} | \Psi \otimes \Psi_{ap} \rangle|^2 \quad (8)$$

be the probability distribution for the final pointer positions. In this expression $|x_f, \mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn}\rangle$ is the simultaneous eigenvector of the Heisenberg-picture operators $\hat{x}_f, \hat{\mu}_{Xf}, \hat{\mu}_{Pf}, \hat{y}_{fj}$, with eigenvalues $x_f, \mu_{Xf}, \mu_{Pf}, y_{fj}$. We continue to employ the notation and terminology of Appleby (1998c). Thus, \hat{x}_f is the final system position operator, $\hat{\mu}_{Xf}$ and $\hat{\mu}_{Pf}$ are the final pointer position operators, and the $\hat{\mu}_{fj}$ represent the additional, internal degrees of freedom characterizing the apparatus.

Let $|x, p\rangle_{\lambda_i} \in \mathcal{H}_{sy}$ be the coherent state with wave function

$$\langle x' | \langle x, p \rangle_{\lambda_i} \rangle = \left(\frac{1}{\pi \lambda_i^2} \right)^{1/4} \exp \left[-\frac{1}{2\lambda_i^2} (x' - x)^2 + \frac{i}{\hbar} px' - \frac{i}{2\hbar} px \right] \quad (9)$$

and let

$$Q_{\lambda_i}(x, p) = \frac{1}{h} |\langle \langle x, p \rangle_{\lambda_i} | \Psi \rangle|^2 \quad (10)$$

be the initial system state Husimi function (Husimi, 1940; Hillery *et al.*, 1984; Lee, 1995; Leonhardt, 1997). We want to show

$$\rho(\mu_{Xf}, \mu_{Pf}) = Q_{\lambda_i}(\mu_{Xf}, \mu_{Pf})$$

for almost all μ_{Xf}, μ_{Pf} whenever the measurement is retrodictively optimal at spatial resolution λ_i (“almost all” being defined relative to ordinary Lebesgue measure on the plane). Our strategy will be to begin by showing that the two functions have the same moments:

$$\int d\mu_{Xf} d\mu_{Pf} \mu_{Xf}^n \mu_{Pf}^m \rho(\mu_{Xf}, \mu_{Pf}) = \int d\mu_{Xf} d\mu_{Pf} \mu_{Xf}^n \mu_{Pf}^m Q_{\lambda_i}(\mu_{Xf}, \mu_{Pf})$$

for every pair of nonnegative integers n, m . Unfortunately, we then face the difficulty that although ρ and Q_{λ_i} are always defined, whatever the initial state of the system, the same is not true of their moments. This is because $\hat{x}_i, \hat{p}_i, \hat{\mu}_{Xf}, \hat{\mu}_{Pf}$ are unbounded operators. The way in which we will circumvent

the difficulty is, first, to prove the result on the assumption that $|\psi\rangle$ is in an appropriately chosen dense subspace of \mathcal{H}_{sy} , and then to use a continuity argument to extend it to the case of arbitrary $|\psi\rangle$.

Let \hat{a}_{λ_i} , $\hat{a}_{\lambda_i}^\dagger$ be the ladder operators

$$\begin{aligned}\hat{a}_{\lambda_i} &= \frac{1}{\sqrt{2}} \left(\frac{1}{\lambda_i} \hat{x}_i + \frac{\lambda_i}{\hbar} \hat{p}_i \right) \\ \hat{a}_{\lambda_i}^\dagger &= \frac{1}{\sqrt{2}} \left(\frac{1}{\lambda_i} \hat{x}_i - \frac{\lambda_i}{\hbar} \hat{p}_i \right)\end{aligned}\quad (11)$$

and define number states $|n\rangle_{\lambda_i} \in \mathcal{H}_{\text{sy}}$ in the usual way, by the requirements

$$\hat{a}_{\lambda_i}|0\rangle_{\lambda_i} = 0, \quad \lambda_i\langle 0|0\rangle_{\lambda_i} = 1, \quad |n\rangle_{\lambda_i} = \frac{1}{\sqrt{n!}} \hat{a}_{\lambda_i}^{\dagger n}|0\rangle_{\lambda_i}$$

(with a slight abuse of notation we sometimes regard the operators \hat{x}_i and \hat{p}_i as acting on \mathcal{H}_{sy} , and sometimes as acting on $\mathcal{H}_{\text{sy}} \otimes \mathcal{H}_{\text{ap}}$). We then define \mathcal{F}_{λ_i} to be the dense subspace of \mathcal{H}_{sy} consisting of all *finite* linear combinations of the vectors $|n\rangle_{\lambda_i}$.

It is easily seen that \mathcal{F}_{λ_i} is in the domain of definition of every polynomial $f(\hat{x}_i, \hat{p}_i)$. In particular, the integral

$$\int dx dp x^n p^m Q_{\lambda_i}(x, p)$$

is defined and finite for all n, m whenever Q_{λ_i} is the Husimi function corresponding to a state in \mathcal{F}_{λ_i} .

Now define the operators

$$\begin{aligned}\hat{b}_{\lambda_i} &= \frac{1}{\sqrt{2}} \left(\frac{1}{\lambda_i} \hat{\mu}_{\text{Xf}} + \frac{i\lambda_i}{\hbar} \hat{\mu}_{\text{Pf}} \right) \\ \hat{b}_{\lambda_i}^\dagger &= \frac{1}{\sqrt{2}} \left(\frac{1}{\lambda_i} \hat{\mu}_{\text{Xf}} - \frac{i\lambda_i}{\hbar} \hat{\mu}_{\text{Pf}} \right)\end{aligned}$$

These operators commute, and so they are certainly not ladder operators. We have

$$\hat{b}_{\lambda_i}^\dagger = \hat{a}_{\lambda_i}^\dagger + \hat{c}_{\lambda_i} \quad (12)$$

where \hat{c}_{λ_i} and $\hat{a}_{\lambda_i}^\dagger$ are the operators defined in Eqs. (6) and (11), respectively. Let $|\psi\rangle$ be any vector $\in \mathcal{F}_{\lambda_i}$. Then $|\psi \otimes \phi_{\text{ap}}\rangle$ is in the domain of $\hat{a}_{\lambda_i}^\dagger$. It is also in the domain of \hat{c}_{λ_i} (the definition of a retrodictively optimal process tacitly assumes that $|\psi \otimes \phi_{\text{ap}}\rangle$ is in the domain of $\hat{\epsilon}_{\text{Xi}}$, $\hat{\epsilon}_{\text{Pi}}$, and therefore in

the domain of \hat{c}_{λ_i} , for all $|\psi\rangle$). It is consequently in the domain of $\hat{b}_{\lambda_i}^\dagger$. Moreover, in view of Lemma 2,

$$\hat{b}_{\lambda_i}^\dagger |\psi \otimes \phi_{\text{ap}}\rangle = (\hat{a}_{\lambda_i}^\dagger |\psi\rangle) \otimes |\phi_{\text{ap}}\rangle$$

where $\hat{a}_{\lambda_i}^\dagger |\psi\rangle$ also $\in \mathcal{F}_{\lambda_i}$. Iterating the argument we conclude that $|\psi \otimes \phi_{\text{ap}}\rangle$ is in the domain of $\hat{b}_{\lambda_i}^{\dagger n}$ and

$$\hat{b}_{\lambda_i}^{\dagger n} |\psi \otimes \phi_{\text{ap}}\rangle = (\hat{a}_{\lambda_i}^{\dagger n} |\psi\rangle) \otimes |\phi_{\text{ap}}\rangle$$

for every nonnegative integer n . Taking adjoints gives

$$\langle \psi \otimes \phi_{\text{ap}} | \hat{b}_{\lambda_i}^m = (\langle \psi | \hat{a}_{\lambda_i}^m \rangle \otimes \langle \phi_{\text{ap}} |$$

for all m . Consequently,

$$\langle \psi \otimes \phi_{\text{ap}} | \hat{b}_{\lambda_i}^m \hat{b}_{\lambda_i}^{\dagger n} | \psi \otimes \phi_{\text{ap}} \rangle = \langle \psi | \hat{a}_{\lambda_i}^m \hat{a}_{\lambda_i}^{\dagger n} | \psi \rangle$$

Now

$$\langle \psi \otimes \phi_{\text{ap}} | \hat{b}_{\lambda_i}^m \hat{b}_{\lambda_i}^{\dagger n} | \psi \otimes \phi_{\text{ap}} \rangle = \int d\mu_{\text{Xf}} d\mu_{\text{Pf}} z_{\lambda_i}^m z_{\lambda_i}^{*n} \rho(\mu_{\text{Xf}}, \mu_{\text{Pf}})$$

where ρ is the distribution of final pointer positions, as defined in Eq. (8), and z_{λ_i} is the complex coordinate

$$z_{\lambda_i} = \frac{1}{\sqrt{2}} \left(\frac{1}{\lambda_i} \mu_{\text{Xf}} + \frac{i\lambda_i}{\hbar} \mu_{\text{Pf}} \right) \tag{13}$$

Also (Hillery *et al.*, 1984; Lee, 1995; Leonhardt, 1997)

$$\langle \psi | \hat{a}_{\lambda_i}^m \hat{a}_{\lambda_i}^{\dagger n} | \psi \rangle = \int d\mu_{\text{Xf}} d\mu_{\text{Pf}} z_{\lambda_i}^m z_{\lambda_i}^{*n} \mathcal{Q}_{\lambda_i}(\mu_{\text{Xf}}, \mu_{\text{Pf}}) \tag{14}$$

where \mathcal{Q}_{λ_i} is the initial-system-state Husimi function, as defined in Eq. (10). Therefore

$$\int d\mu_{\text{Xf}} d\mu_{\text{Pf}} z_{\lambda_i}^m z_{\lambda_i}^{*n} \rho(\mu_{\text{Xf}}, \mu_{\text{Pf}}) = \int d\mu_{\text{Xf}} d\mu_{\text{Pf}} z_{\lambda_i}^m z_{\lambda_i}^{*n} \mathcal{Q}_{\lambda_i}(\mu_{\text{Xf}}, \mu_{\text{Pf}})$$

for all n, m . It follows that

$$\int d\mu_{\text{Xf}} d\mu_{\text{Pf}} f(z_{\lambda_i}, z_{\lambda_i}^*) \rho(\mu_{\text{Xf}}, \mu_{\text{Pf}}) = \int d\mu_{\text{Xf}} d\mu_{\text{Pf}} f(z_{\lambda_i}, z_{\lambda_i}^*) \mathcal{Q}_{\lambda_i}(\mu_{\text{Xf}}, \mu_{\text{Pf}})$$

for every polynomial f . In particular,

$$\begin{aligned} & \int d\mu_{Xf} d\mu_{Pf} \mu_{Xf}^m \mu_{Pf}^n \rho(\mu_{Xf}, \mu_{Pf}) \\ &= \int d\mu_{Xf} d\mu_{Pf} \mu_{Xf}^m \mu_{Pf}^n Q_{\lambda_i}(\mu_{Xf}, \mu_{Pf}) \end{aligned} \quad (15)$$

for all m, n .

At this stage one needs to be careful. It is tempting to suppose that two probability measures which have the same moments must be equal. In fact, this inference is not always justified (see Reed and Simon, 1975, Vol. 2). However, it is justified here, as we show in the Appendix. Consequently

$$\rho(\mu_{Xf}, \mu_{Pf}) = Q_{\lambda_i}(\mu_{Xf}, \mu_{Pf}) \quad (16)$$

for almost all μ_{Xf}, μ_{Pf} whenever the initial system state $|\psi\rangle$ is in the space \mathcal{F}_{λ_i} .

It remains for us to show that the distributions are equal in the case of arbitrary $|\psi\rangle \in \mathcal{H}_{\text{sy}}$. We will do this by using a continuity argument.

Choose a sequence $|\psi_n\rangle \in \mathcal{F}_{\lambda_i}$ converging to $|\psi\rangle$. Let $Q_{\lambda_i, n}$ be the Husimi function and ρ_n the distribution of measured values corresponding to $|\psi_n\rangle$. Let Q_{λ_i} be the Husimi function and ρ the distribution of measured values corresponding to $|\psi\rangle$.

We have, as an immediate consequence of the definition, Eq. (10),

$$Q_{\lambda_i}(\mu_{Xf}, \mu_{Pf}) = \lim_{n \rightarrow \infty} (Q_{\lambda_i, n}(\mu_{Xf}, \mu_{Pf})) \quad (17)$$

for all μ_{Xf}, μ_{Pf} .

On the other hand, it is not generally true that ρ_n converges pointwise to ρ . It does, however, contain a subsequence which converges pointwise almost everywhere. In fact, let \mathcal{L}_1 be the Banach space consisting of all integrable functions on \mathbb{R}^2 with norm

$$\|f\|_1 = \int d\mu_{Xf} d\mu_{Pf} |f(\mu_{Xf}, \mu_{Pf})|$$

We have

$$\begin{aligned} \|\rho - \rho_n\|_1 &= \int d\mu_{Xf} d\mu_{Pf} \left| \int dx_f dy_{f1} \dots y_{fn} (|\langle x_f, \mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn} | \psi \right. \\ &\quad \left. \otimes \phi_{\text{ap}} \rangle|^2 - |\langle x_f, \mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn} | \psi_n \otimes \phi_{\text{ap}} \rangle|^2) \right| \\ &\leq \| |\psi \otimes \phi_{\text{ap}} \rangle - |\psi_n \otimes \phi_{\text{ap}} \rangle \| (\| |\psi \otimes \phi_{\text{ap}} \rangle \| + \| |\psi_n \otimes \phi_{\text{ap}} \rangle \|) \\ &\rightarrow 0 \end{aligned}$$

We see from this that $\rho_n \rightarrow \rho$ in the topology of \mathcal{L}_1 . We may therefore use the Riesz–Fisher theorem (Reed and Simon, 1975, Vol. 1) to deduce that it contains a subsequence ρ_{n_r} such that

$$\rho(\mu_{Xf}, \mu_{Pf}) = \lim_{r \rightarrow \infty} (\rho_{n_r}(\mu_{Xf}, \mu_{Pf}))$$

for almost all μ_{Xf}, μ_{Pf} . In view of this result, Eq. (17), and the fact that

$$\rho_{n_r}(\mu_{Xf}, \mu_{Pf}) = \Omega_{\lambda_i n_r}(\mu_{Xf}, \mu_{Pf})$$

for all r and almost all μ_{Xf}, μ_{Pf} we deduce that

$$\rho(\mu_{Xf}, \mu_{Pf}) = \Omega_{\lambda_i}(\mu_{Xf}, \mu_{Pf})$$

for almost all μ_{Xf}, μ_{Pf} .

3. PREDICTIVELY OPTIMAL MEASUREMENTS

We will say that a simultaneous measurement process of the kind defined in Appleby (1998c) is *predictively optimal* if the product of predictive errors is minimized:

$$\Delta_{eF} \Delta_{eP} = \frac{\hbar}{2} \tag{18}$$

In view of the commutation relation

$$[\hat{\epsilon}_{Xf}, \hat{\epsilon}_{Pf}] = i\hbar \tag{19}$$

there is no need to impose the condition that the measurement be predictively unbiased as a separate requirement: it is a consequence of the condition of Eq. (18).

Equations (18) and (19) together imply

$$\langle \psi \otimes \phi_{ap} | \hat{\epsilon}_{Xf}^2 | \psi \otimes \phi_{ap} \rangle \langle \psi \otimes \phi_{ap} | \hat{\epsilon}_{Pf}^2 | \psi \otimes \phi_{ap} \rangle = \frac{\hbar^2}{4}$$

for every normalized $|\psi\rangle \in \mathcal{H}_{sy}$. By an argument which parallels the proof of Lemma 1 we infer that there exists a fixed number λ_f such that

$$\langle \psi \otimes \phi_{ap} | \hat{\epsilon}_{Xf}^2 | \psi \otimes \phi_{ap} \rangle = \frac{\lambda_f^2}{2}$$

$$\langle \psi \otimes \phi_{ap} | \hat{\epsilon}_{Pf}^2 | \psi \otimes \phi_{ap} \rangle = \frac{\hbar^2}{2\lambda_f^2}$$

for every normalized $|\psi\rangle \in \mathcal{H}_{\text{sy}}$. It is then straightforward to show that

$$\hat{d}_{\lambda_f} |\psi \otimes \phi_{\text{ap}}\rangle = 0 \quad (20)$$

for all $|\psi\rangle \in \mathcal{H}_{\text{sy}}$, where \hat{d}_{λ_f} is the annihilation operator

$$\hat{d}_{\lambda_f} = \frac{1}{\sqrt{2}} \left(\frac{1}{\lambda_f} \hat{\epsilon}_{\text{Xf}} + \frac{i\lambda_f}{\hbar} \hat{\epsilon}_{\text{Pf}} \right)$$

Since $\hat{\epsilon}_{\text{Xf}}$, $\hat{\epsilon}_{\text{Pf}}$ are canonically conjugate, there exist kets $|\epsilon_{\text{Xf}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{f1}, \dots, y_{fn}\rangle_{\epsilon}$ which are simultaneous eigenvectors of the operators $\hat{\epsilon}_{\text{Xf}}$, $\hat{\mu}_{\text{Xf}}$, $\hat{\mu}_{\text{Pf}}$, \hat{y}_{fj} , and which have the property

$$\begin{aligned} & {}_{\epsilon}\langle \epsilon_{\text{Xf}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{f1}, \dots, y_{fn} | \hat{\epsilon}_{\text{Pf}} | \Psi \rangle \\ &= -i\hbar \frac{\partial}{\partial \epsilon_{\text{Xf}}} {}_{\epsilon}\langle \epsilon_{\text{Xf}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{f1}, \dots, y_{fn} | \Psi \rangle \end{aligned} \quad (21)$$

for all $|\Psi\rangle \in \mathcal{H}_{\text{sy}} \otimes \mathcal{H}_{\text{ap}}$. In view of Eq. (20) we then have

$$\left(\frac{1}{\lambda_f} \epsilon_{\text{Xf}} + \lambda_f \frac{\partial}{\partial \epsilon_{\text{Xf}}} \right) {}_{\epsilon}\langle \epsilon_{\text{Xf}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{f1}, \dots, y_{fn} | \psi \otimes \phi_{\text{ap}} \rangle = 0$$

for all $|\psi\rangle \in \mathcal{H}_{\text{sy}}$. Solving this equation, we find

$$\begin{aligned} & {}_{\epsilon}\langle \epsilon_{\text{Xf}}, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{f1}, \dots, y_{fn} | \psi \otimes \phi_{\text{ap}} \rangle \\ &= \left(\frac{1}{\pi \lambda_f^2} \right)^{1/4} \exp \left[-\frac{1}{2\lambda_f^2} \epsilon_{\text{Xf}}^2 \right] \Phi(\mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{f1}, \dots, y_{fn}) \end{aligned} \quad (22)$$

where Φ is an arbitrary normalised function.

There also exist kets $|x_f, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{f1}, \dots, y_{fn}\rangle_x$ which are simultaneous eigenvectors of the operators \hat{x}_f , $\hat{\mu}_{\text{Xf}}$, $\hat{\mu}_{\text{Pf}}$, \hat{y}_{fj} with the property

$$\begin{aligned} & {}_x\langle x_f, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{f1}, \dots, y_{fn} | \hat{p}_f | \Psi \rangle \\ &= -i\hbar \frac{\partial}{\partial x_f} {}_x\langle x_f, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{f1}, \dots, y_{fn} | \Psi \rangle \end{aligned} \quad (23)$$

for all $|\Psi\rangle \in \mathcal{H}_{\text{sy}} \otimes \mathcal{H}_{\text{ap}}$. In view of the defining relation $\hat{\epsilon}_{\text{Xf}} = \hat{\mu}_{\text{Xf}} - \hat{x}_f$ we must have

$$\begin{aligned} & |x_f, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{f1}, \dots, y_{fn}\rangle_x \\ &= e^{-i\chi(x_f, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{f1}, \dots, y_{fn})} |\mu_{\text{Xf}} - x_f, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{f1}, \dots, y_{fn}\rangle_{\epsilon} \end{aligned} \quad (24)$$

where $\exp[i\chi(x_f, \mu_{\text{Xf}}, \mu_{\text{Pf}}, y_{f1}, \dots, y_{fn})]$ is a phase. In view of Eqs. (21) and (23) we must then have

$$\begin{aligned} & \langle x_f | \mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn} | \hat{p}_f | \Psi \rangle \\ &= -i\hbar \frac{\partial}{\partial x_f} \left(e^{i\chi(x_f, \mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn})} \langle \mu_{Xf} - x_f, \mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn} | \Psi \rangle \right) \end{aligned}$$

and

$$\begin{aligned} & \langle x_f, \mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn} | \hat{p}_f | \Psi \rangle \\ &= e^{i\chi(x_f, \mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn})} \langle \mu_{Xf} - x_f, \mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn} | \hat{\mu}_{Pf} - \hat{\epsilon}_{Pf} | \Psi \rangle \\ &= e^{i\chi(x_f, \mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn})} \\ & \quad \times \left(\mu_{Pf} + i\hbar \frac{\partial}{\partial \hat{\epsilon}_{Xf}} \right) \langle \hat{\epsilon}_{Xf}, \mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn} | \Psi \rangle |_{\hat{\epsilon}_{Xf} = \mu_{Xf} - x_f} \end{aligned}$$

for all $|\Psi\rangle \in \mathcal{H}_{sy} \otimes \mathcal{H}_{ap}$. Hence

$$\hbar \frac{\partial}{\partial x_f} \chi(x_f, \mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn}) = \mu_{Pf}$$

which implies

$$\chi(x_f, \mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn}) = \frac{1}{\hbar} \mu_{Pf} x_f + \chi_0(\mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn})$$

where χ_0 is an arbitrary function. Using this result and Eq. (24) in Eq. (22), we deduce that the final-state wave function can be written

$$\begin{aligned} & \langle x_f, \mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn} | \psi \otimes \phi_{ap} \rangle \\ &= \left(\frac{1}{\pi \lambda_f^2} \right)^{1/4} \exp \left[-\frac{1}{2\lambda_f^2} (\mu_{Xf} - x_f)^2 + \frac{i}{\hbar} \mu_{Pf} x_f + i\chi_0(\mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn}) \right] \\ & \quad \times \Phi(\mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn}) \end{aligned}$$

In terms of the state $|\mu_{Xf}, \mu_{Pf}\rangle_{\lambda_f}$ defined in Eq. (9) this becomes

$$\begin{aligned} & \langle x_f, \mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn} | \psi \otimes \phi_{ap} \rangle \\ &= \langle x_f | \mu_{Xf}, \mu_{Pf} \rangle_{\lambda_f} \Phi'(\mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn}) \end{aligned}$$

where

$$\begin{aligned} & \Phi'(\mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn}) \\ &= \exp \left[i\chi_0(\mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn}) + \frac{i}{2\hbar} \mu_{Pf} \mu_{Xf} \right] \\ & \quad \times \Phi(\mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn}) \end{aligned}$$

The distribution of measured values $\rho(\mu_{Xf}, \mu_{Pf})$ can be written in terms of Φ' :

$$\rho(\mu_{Xf}, \mu_{Pf}) = \int dy_{f1} \dots dy_{fn} |\Phi'(\mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn})|^2$$

Suppose, now, that the pointer positions are found to be in the region $\mathcal{R} \subseteq \mathbb{R}^2$. Let $\hat{\rho}_{\text{sy}}$ be the reduced density matrix describing the state of the system immediately afterward. Then

$$\begin{aligned} \langle X_{f1} | \hat{\rho}_{\text{sy}} | X_{f2} \rangle &= \frac{1}{p_{\mathcal{R}}} \int_{\mathcal{R} \times \mathbb{R}^n} d\mu_{Xf} d\mu_{Pf} dy_{f1} \dots dy_{fn} |\Phi'(\mu_{Xf}, \mu_{Pf}, y_{f1}, \dots, y_{fn})|^2 \\ & \quad \times \langle X_{f1} | (\mu_{Xf}, \mu_{Pf})_{\lambda_f} \rangle \langle (\mu_{Xf}, \mu_{Pf})_{\lambda_f} | X_{f2} \rangle \end{aligned}$$

where $p_{\mathcal{R}}$ is the probability of finding $(\mu_{Xf}, \mu_{Pf}) \in \mathcal{R}$:

$$p_{\mathcal{R}} = \int_{\mathcal{R}} d\mu_{Xf} d\mu_{Pf} \rho(\mu_{Xf}, \mu_{Pf})$$

Hence

$$\hat{\rho}_{\text{sy}} = \frac{1}{p_{\mathcal{R}}} \int_{\mathcal{R}} d\mu_{Xf} d\mu_{Pf} \rho(\mu_{Xf}, \mu_{Pf}) |(\mu_{Xf}, \mu_{Pf})_{\lambda_f}\rangle \langle (\mu_{Xf}, \mu_{Pf})_{\lambda_f}|$$

On the other hand,

$$\hat{\rho}_{\text{sy}} = \int d\mu_{Xf} d\mu_{Pf} P_{\lambda_f}(\mu_{Xf}, \mu_{Pf}) |(\mu_{Xf}, \mu_{Pf})_{\lambda_f}\rangle \langle (\mu_{Xf}, \mu_{Pf})_{\lambda_f}|$$

where P_{λ_f} is the anti-Husimi function, or P -function (Glauber, 1963; Sudarsan, 1963; Hillery *et al.*, 1984; Lee, 1995; Leonhardt, 1997) describing the final state of the system. Comparing these expressions, we see that

$$P_{\lambda_f}(\mu_{Xf}, \mu_{Pf}) = \begin{cases} (1/p_{\mathcal{R}})\rho(\mu_{Xf}, \mu_{Pf}) & \text{if } (\mu_{Xf}, \mu_{Pf}) \in \mathcal{R} \\ 0 & \text{otherwise} \end{cases} \quad (25)$$

If \mathcal{R} is a sufficiently small region centered on the point (μ_{Xf}, μ_{Pf}) , the system is approximately in the state $|(\mu_{Xf}, \mu_{Pf})_{\lambda_f}\rangle$ after the measurement:

$$\hat{\rho}_{\text{sy}} \approx |(\mu_{\text{Xf}}, \mu_{\text{Pf}})_{\lambda_f}\rangle\langle(\mu_{\text{Xf}}, \mu_{\text{Pf}})_{\lambda_f}|$$

Equation (25) shows that the effect of a predictively optimal measurement process is to leave the system in a state for which P_{λ_f} is a probability density function. Such states are, of course, exceptional. In many cases, P_{λ_f} is not even defined as a tempered distribution (Hillery *et al.*, 1984; Lee, 1995; Leonhardt, 1997).

4. THE INTERPRETATION OF THE HUSIMI FUNCTION

The result proved in Section 2 shows that there is a certain analogy between the Husimi function and the x -space probability density function $|\langle x|\psi\rangle|^2$. To see this, let us examine just what is meant by the statement that $|\langle x|\psi\rangle|^2 \delta x$ represents the probability of finding the position to lie in the interval $(x, x + \delta x)$.

Consider a measurement of x only. For the sake of simplicity suppose that the measuring apparatus has only one degree of freedom, corresponding to the single pointer observable $\hat{\mu}_X$ (the argument which follows does not depend on this assumption, however). Let $|\psi\rangle$ and $|\phi_{\text{ap}}\rangle$ be the initial states of the system and apparatus, respectively, and let \hat{U} be the unitary evolution operator describing the measurement interaction. Let $\hat{x}_i = \hat{x}$ and $\hat{\mu}_{\text{Xf}} = \hat{U}^\dagger \mu_{\text{Xf}} \hat{U}$ be the Heisenberg-picture operators describing the initial position of the system and final position of the pointer, respectively. Let $\hat{\epsilon}_{\text{Xi}} = \hat{\mu}_{\text{Xf}} - \hat{x}_i$ be the retrodictive error operator.

The final-state wave function can be written (in the Schrödinger picture)

$$\langle x, \mu_X | \hat{U} | \psi \otimes \phi_{\text{ap}} \rangle = \int dx' K(x, \mu_X; x') \langle x' | \psi \rangle$$

for some kernel K . The probability distribution describing the result of the measurement then takes the form

$$\rho(\mu_X) = \int dx | \int dx' K(x, \mu_X; x') \langle x' | \psi \rangle |^2 \tag{26}$$

After a certain amount of algebra one also finds

$$\langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{\text{Xi}}^2 | \psi \otimes \phi_{\text{ap}} \rangle = \int dx d\mu_X | \int dx' (\mu_X - x') K(x, \mu_X; x') \langle x' | \psi \rangle |^2 \tag{27}$$

Suppose that $\Delta_{\epsilon;X} = 0$. Then we see from Eq. (27) that K must take the form

$$K(x, \mu_X; x') = f(x, \mu_X) \delta(\mu_X - x')$$

for some function f . The unitarity of \hat{U} means that f must satisfy

$$\int dx |f(x, \mu_x)|^2 = 1$$

Using these results in Eq. (26), we find

$$\rho(\mu_x) = |\langle \mu_x | \psi \rangle|^2$$

whenever the measurement is perfectly accurate for the purposes of retrodiction.

Suppose, on the other hand, that $\Delta_{e_i} x > 0$. Then $\rho(\mu_x)$ will not generally coincide with the function $|\langle \mu_x | \psi \rangle|^2$. If $\Delta_{e_i} x$ is small compared with the de Broglie wavelength, then we see from Eqs. (26) and (27) that $\rho(\mu_x) \approx |\langle \mu_x | \psi \rangle|^2$. Otherwise, we do not expect the two functions even to be approximately equal.

Although one may possibly approach, one does not expect actually to achieve the limit of perfect accuracy. It follows that one does not expect the function $|\langle \mu_x | \psi \rangle|^2$ to describe the outcome of any practically realizable measurement of position.

This being so, what, exactly, is the significance of the function $|\langle \mu_x | \psi \rangle|^2$? In the first place, it serves as a standard of comparison against which the outcome of experimentally realizable measurements can be judged, in the sense that the better the measurement, the more closely does the function $|\langle \mu_x | \psi \rangle|^2$ approximate the distribution of actual results.

In the second place, we see from Eq. (26) that the outcome of a real measurement of position depends, not only on the state of the system, via the function $\langle x' | \psi \rangle$, but also on the details of the measurement process, via the function $K(x, \mu_x; x')$. In the limit of perfect retrodictive accuracy, however, the dependence on the apparatus (as represented by the kernel K) disappears, and the distribution of results is determined solely by the state of the system (as represented by the vector $|\psi\rangle$). $|\langle \mu_x | \psi \rangle|^2$ does, so to speak, represent the *intrinsic* distribution of position, independent of any properties specific to the particular measuring instrument employed. In a real measurement, by contrast, the outcome is (in a manner of speaking) contaminated by instrumental contributions, which one may try to reduce, but can never entirely eliminate.

One typically regards the function $|\langle \mu_x | \psi \rangle|^2$ simply, and without qualification, as *the* x -space probability distribution. It owes this canonical status to the two features just mentioned. The result proved in Section 2 shows that the Husimi function has analogous features. It describes the outcome of those measurements which are retrodictively optimal, or “best.” It is otherwise independent of the details of the particular process considered. It might

therefore be regarded as the canonical probability distribution for position and momentum.

In classical mechanics one has the concept of the “actual” distribution describing an ensemble of identically prepared systems. Quantum mechanics contains no precise analogue for this concept [unless one adopts a “hidden-variables” interpretation, as discussed by Bohm and Hiley (1993) or Holland (1993), for example]. Nevertheless, the result proved in Section 2 shows that there are certain resemblances between the Husimi function and the classical distribution. The Husimi function is clearly not the same as the classical distribution. However, one might reasonably argue that it is the closest that quantum mechanics allows us to get to the concept of a “real” or “objective” phase space probability distribution.

APPENDIX. PROOF OF EQUATION (16)

Rather than working in terms of the functions ρ, Q_{λ_i} it will be convenient, instead, to work in terms of the measures

$$d\mu_\rho = \rho(\mu_{Xf}, \mu_{Pf}) d\mu_{Xf} d\mu_{Pf}$$

$$d\mu_Q = Q_{\lambda_i}(\mu_{Xf}, \mu_{Pf}) d\mu_{Xf} d\mu_{Pf}$$

We have from Eqs. (14) and (15)

$$\int d\mu_\rho |z_{\lambda_i}|^{2n} = \int d\mu_Q |z_{\lambda_i}|^{2n} = \langle \psi | \hat{a}_{\lambda_i}^n \hat{a}_{\lambda_i}^{\dagger n} | \psi \rangle \tag{A1}$$

where z_{λ_i} is the complex coordinate defined in Eq. (13). Our strategy will be, first, to establish a bound on the rate at which these quantities grow with increasing n , and then to use this to show that the measures μ_ρ, μ_Q have the same Fourier transform.

$|\psi\rangle$ is in the subspace \mathcal{F}_{λ_i} . It can therefore be written

$$|\psi\rangle = \sum_{r=0}^l c_r |r\rangle_{\lambda_i}$$

for some integer l . Hence

$$\langle \psi | \hat{a}_{\lambda_i}^n \hat{a}_{\lambda_i}^{\dagger n} | \psi \rangle = \sum_{r=0}^l \frac{(n+r)!}{r!} |c_r|^2 \leq \frac{(n+l)!}{l!}$$

Let μ stand for either of the measures μ_ρ, μ_Q . In view of the inequality just proved, Eq. (A1), and the fact

$$|z_{\lambda_i}|^{2n+1} \leq \frac{1}{2} (|z_{\lambda_i}|^{2n} + |z_{\lambda_i}|^{2n+2})$$

we have

$$\int d\mu |z_{\lambda_i}|^n \leq \frac{\Gamma(\frac{1}{3}n + l + 3/2)}{\Gamma(l + 1)}$$

for every nonnegative integer n . Hence

$$\sum_{n=0}^{\infty} \frac{1}{n!} \int d\mu |\beta z_{\lambda_i} + \gamma z_{\lambda_i}^*|^n < \infty$$

for all $\beta, \gamma \in \mathbb{C}$. It follows that the functions $\exp|\beta z_{\lambda_i} + \gamma z_{\lambda_i}^*|$ and $\exp(\beta z_{\lambda_i} + \gamma z_{\lambda_i}^*)$ are μ -integrable. We may therefore use Lebesgue's dominated convergence theorem (Reed and Simon, 1975, Vol. 1) to infer

$$\begin{aligned} \int d\mu_{\rho} \exp[\beta z_{\lambda_i} + \gamma z_{\lambda_i}^*] &= \lim_{N \rightarrow \infty} \left(\sum_{n=0}^N \frac{1}{n!} \int d\mu_{\rho} (\beta z_{\lambda_i} + \gamma z_{\lambda_i}^*)^n \right) \\ &= \lim_{N \rightarrow \infty} \left(\sum_{n=0}^N \frac{1}{n!} \int d\mu_{\mathcal{Q}} (\beta z_{\lambda_i} + \gamma z_{\lambda_i}^*)^n \right) \\ &= \int d\mu_{\mathcal{Q}} \exp[\beta z_{\lambda_i} + \gamma z_{\lambda_i}^*] \end{aligned}$$

for all $\beta, \gamma \in \mathbb{C}$. Consequently,

$$\int d\mu_{\rho} \exp[i(k_X \mu_{Xf} + k_P \mu_{Pf})] = \int d\mu_{\mathcal{Q}} \exp[i(k_X \mu_{Xf} + k_P \mu_{Pf})]$$

for all $k_X, k_P \in \mathbb{R}$. Inverting the Fourier transforms, we deduce

$$\mu_{\rho} = \mu_{\mathcal{Q}}$$

REFERENCES

- Ali, S. T., and Prugovečki, E. (1977). *Journal of Mathematical Physics*, **18**, 219.
 Appleby, D. M. (1998a). *International Journal of Theoretical Physics*, **37**, 1491.
 Appleby, D. M. (1998b). *Journal of Physics A*, **31**, 6419.
 Appleby, D. M. (1998c). *International Journal of Theoretical Physics*, **37**, 2557.
 Arthurs, E., and Kelly, J. L. (1965). *Bell System Technical Journal*, **44**, 725.
 Ban, M. (1997). *International Journal of Theoretical Physics*, **36**, 2583.
 Banaszek, K., and Wódkiewicz, K. (1997). *Physical Review A*, **55**, 3117.
 Bohm, D., and Hiley, B. J. (1993). *The Undivided Universe*, Routledge, London.

- 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.
- Burak, D., and Wódkiewicz, K. (1992). *Physical Review A*, **46**, 2744.
- Busch, P. (1985). *International Journal of Theoretical Physics*, **24**, 63.
- Busch, P., Grabowski, M., and Lahti, P. J. (1995). *Operational Quantum Physics*, Springer-Verlag, Berlin.
- Englert, B.-G., and Wódkiewicz, K. (1995). *Physical Review A*, **51**, R2661.
- Glauber, R. J. (1963). *Physical Review*, **131**, 2766.
- Halliwell, J. J. (1992). *Physical Review D*, **46**, 1610.
- Hillery, M., O'Connell, R. F., Scully, M. O., and Wigner, E. P. (1984). *Physics Reports*, **106**, 121.
- Holevo, A. S. (1982). *Probabilistic and Statistical Aspects of Quantum Theory*, North-Holland, Amsterdam.
- Holland, P. R. (1993). *The Quantum Theory of Motion*, Cambridge University Press, Cambridge.
- Husimi, K. (1940). *Proceedings of the Physical and Mathematical Society of Japan*, **22**, 264.
- Lee, H. W. (1995). *Physics Reports*, **259**, 147.
- Leonhardt, U. (1997). *Measuring the Quantum State of Light*, Cambridge University Press, Cambridge.
- Leonhardt, U., and Paul, H. (1993). *Journal of Modern Optics*, **40**, 1745.
- Leonhardt, U., and Paul, H. (1995). *Progress in Quantum Electronics*, **19**, 89.
- Marchiolli, M. A., Mizrahi, S. S., and Dodonov, V. V. (1997). *Physical Review A*, **56**, 4278.
- Power, W. L., Tan, S. M., and Wilkens, M. (1997). *Journal of Modern Optics*, **44**, 2591.
- 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 Space-time*, Reidel, Dordrecht.
- Reed, M., and Simon, B. (1975). *Methods of Modern Mathematical Physics*, Vols. 1–4, Academic Press, New York.
- Stenholm, S. (1992). *Annals of Physics*, **218**, 233.
- Sudarshan, E. C. G. (1963). *Physical Review Letters*, **10**, 277.
- 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. (1984). *Physical Review Letters*, **52**, 1064.
- Wódkiewicz, K. (1986). *Physics Letters A*, **115**, 304.
- Wódkiewicz, K. (1987). *Physics Letters A*, **124**, 207.