

Maximal accuracy and minimal disturbance in the Arthurs-Kelly simultaneous measurement process

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1998 J. Phys. A: Math. Gen. 31 6419

(<http://iopscience.iop.org/0305-4470/31/30/009>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.102

The article was downloaded on 02/06/2010 at 07:08

Please note that [terms and conditions apply](#).

Maximal accuracy and minimal disturbance in the Arthurs–Kelly simultaneous measurement process

D M Appleby†

Department of Physics, Queen Mary and Westfield College, Mile End Road, London E1 4NS, UK

Received 24 October 1997, in final form 8 May 1998

Abstract. The accuracy of the Arthurs–Kelly model of a simultaneous measurement of position and momentum is analysed using concepts developed by Braginsky and Khalili in the context of measurements of a single quantum observable. A distinction is made between the errors of retrodiction and prediction. It is shown that the distribution of measured values coincides with the initial-state Husimi function when the retrodictive accuracy is maximized (this being the result first obtained by Busch, using somewhat different methods). It is further shown that the distribution of measured values is related to the final-state anti-Husimi function (the P representation of quantum optics) when the predictive accuracy is maximized. The disturbance of the system by the measurement is also discussed. A class of minimally disturbing measurements is characterized. It is shown that the distribution of measured values then coincides with one of the smoothed Wigner functions described by Cartwright.

1. Introduction

It has been known since the publication of Heisenberg's original paper [1] that quantum mechanics does not allow both the position and momentum of a system to be measured with arbitrary accuracy. However, it does not follow from this that one cannot measure both quantities with a less than perfect degree of accuracy. Indeed, it would seem that it is essential that quantum mechanics does permit such measurements, if it is to be possible to derive classical mechanics from quantum mechanics as an approximate theory, valid in some appropriate limit. For this and other reasons simultaneous measurements have been the subject of much theoretical interest over the years.

Recently interest in them has been greatly increased, owing to technical advances in the field of quantum optics. As a result of these advances simultaneous, imperfect measurements of the position and momentum of a quantum mechanical system are no longer confined to the idealized world of *gedanken* experiments. They can actually be realized in the laboratory [2–9]. For a recent review of these methods, and of the closely related subject of the tomographic reconstruction of the quantum state, the reader may consult Leonhardt and Paul [10], or Leonhardt [11].

In the following we shall be concerned with the problem of describing the accuracy of such measurements, and the disturbance they produce in the system whose position and momentum are being measured. In [12] we developed some general methods for analysing this problem. Our purpose now is to illustrate the use of these methods, by applying them to the example of the Arthurs–Kelly process [2–4, 9, 11].

† E-mail address: D.M.Appleby@qmw.ac.uk

The accuracy of, and disturbance produced by, a simultaneous measurement process has been the subject of many previous investigations. One important relationship is the inequality originally derived by Arthurs and Kelly [2, 6, 11, 13, 14]

$$\Delta\mu_{Xf}\Delta\mu_{Pf} \geq \hbar \quad (1)$$

where $\Delta\mu_{Xf}$ and $\Delta\mu_{Pf}$ denote the uncertainties in the final pointer positions, after the measurement has been completed. This inequality is to be compared with the ordinary uncertainty principle

$$\Delta x \Delta p \geq \frac{\hbar}{2}. \quad (2)$$

The doubling of the lower bound in equation (1) as compared with equation (2) is due to the additional noise introduced by the measurement process.

We discussed the Arthurs–Kelly inequality in [12]. Here we shall content ourselves by remarking that, although the product $\Delta\mu_{Xf}\Delta\mu_{Pf}$ provides a numerical indication of the overall accuracy of the measurement, the quantities $\Delta\mu_{Xf}$ and $\Delta\mu_{Pf}$ are not themselves directly interpretable as experimental errors. One would like to find quantities which can be individually and immediately identified with the errors in the measured values of position and momentum.

The problem of defining quantities which do provide a direct characterization of the errors in a simultaneous measurement of position and momentum has been systematically addressed within the framework of the approach based on the concepts of a positive operator-valued measure (POVM) and an unsharp observable [15–28]. For a recent review, and additional references, the reader may consult the book by Busch *et al* [9]. We give a brief discussion of the way in which the measurement errors are defined in this approach in the appendix.

Our own interest in this problem stems from the work of Uffink [29], who has criticized the POVM approach, and more generally the very concept of a joint measurement of non-commuting observables, in terms which may fairly be described as unequivocal. He states on p 200:

‘... the formalism of quantum theory, as it is presented by von Neumann, simply has no room for a description of a joint measurement of position and momentum at all.’

He further asserts

‘the claim that within this formalism [*i.e.* the formalism based on on the concept of a POVM] a joint unsharp measurement of position and momentum or a pair of spin components is possible is false.’

He goes on to assert, that claims that the POVM approach does provide a theory of joint measurements,

‘rest on the adoption of inappropriate definitions, *i.e.*, definitions that trivialize the problem.’

In connection with these statements, we should say at the outset, that we do not feel that Uffink’s arguments are sufficiently strong to support his conclusion. On a purely intuitive level, it seems evident that the processes discussed in [2–11] must be describable as joint measurements in some sense of the word. It seems no less evident, that it cannot really be correct to dismiss the claims of the POVM approach in the way that Uffink does dismiss

them, wholesale, as ‘false’ and/or ‘trivial’. In other words, it appears to us that Uffink overstates his case. However, that is not to say that his criticisms are without foundation. On the contrary, we consider that Uffink has raised some very important and interesting questions of principle which it would be worthwhile trying to clarify. This was the original motive for the work reported in [12].

If one examines the statements just quoted it will be perceived that Uffink actually makes two different claims, which may be enumerated as follows.

(1) The conventional, generally accepted formulation of quantum mechanics, as it is presented in the book by von Neumann [30], has no room for the concept of a simultaneous measurement of position and momentum. Or, to put it another way, the conventional theory of quantum mechanics does not permit both the position and momentum of a system to be jointly determined, not even with a less than perfect degree of accuracy.

(2) The attempt to solve the joint measurement problem within the framework of the approach based on POVMs and unsharp observables fails, because it depends on definitions which are ‘inappropriate’ because they ‘trivialize’ the problem.

In [12] we addressed the first of these objections (we intend to address the second objection in a future publication). We assumed that a measurement consists of a process in which a correlation is established between one set of observables (the quantities being measured) and another set of observables (the meter readings). In the case of interest to us, the observables being measured are the position \hat{x} and momentum \hat{p} of a quantum mechanical system having one degree of freedom. The observables representing the measurement outcome are two commuting pointer observables $\hat{\mu}_X$ and $\hat{\mu}_P$. A process which brings about, simultaneously, both a correlation between \hat{x} and $\hat{\mu}_X$, and a correlation between \hat{p} and $\hat{\mu}_P$, is *ipso facto* describable as a simultaneous measurement of \hat{x} and \hat{p} . The problem then reduces to the problem of giving a precise characterization of the degree of this correlation. In other words, it reduces to the problem of giving a precise characterization of the measurement accuracy. In [12] we showed how this problem can be solved by taking some concepts developed by Braginsky and Khalili [31] in the context of single measurements of position only, and applying them to the case of simultaneous measurements of position and momentum together. At no stage did we go beyond the conventional theory of quantum mechanics, as it is presented by von Neumann. We were therefore able to conclude, that contrary to what Uffink asserts, the conventional theory does permit the simultaneous, albeit less than perfectly accurate, measurement of both the position and the momentum of a quantum mechanical system.

Our motive for undertaking the research reported in [12] was the desire to clarify a fundamental question of principle. However, in the course of this research it became clear to us that the methods we had developed had an interest which went somewhat beyond the purpose for which we had originally intended them. It appears to us that they provide a way of analysing the physics of simultaneous measurement processes which is useful, both as a mathematical technique, and as a source of physical insight.

We discussed the fundamental concepts on which our approach is based in [12] (see also section 2 below). Our approach rests on making a careful distinction between the retrodictive and predictive (or determinative and preparative) aspects of a joint measurement process. Corresponding to this distinction we introduce two different kinds of error: a pair of retrodictive errors and a pair of predictive errors. We also define a pair of disturbances, providing a quantitative indication of the disturbance of the system by the measurement process. We then derive a total of six inequalities relating these quantities: namely a retrodictive error relationship, a predictive error relationship, and four error-disturbance relationships. As we stated above, our approach represents a development of the analysis

given by Braginsky and Khalili [31] for the case of single measurements of position only.

In the case of joint measurement processes, Ali and Prugovečki [18] gave a definition which is essentially equivalent to our definition of the errors of retrodiction for the class of processes to which their definition is applicable. In particular, our definition is essentially equivalent to theirs in the case of the Arthurs–Kelly process, which forms the subject of this paper. However, it can be shown (see the appendix) that our definition is the more general of the two, in that it applies to a class of measurement processes which is strictly larger than the class considered by Ali and Prugovečki. Also, our definition provides additional physical insight even in those cases where it is equivalent to—gives the same numerical values as—the definition of Prugovečki and Ali. We used this fact in [12], when trying to answer Uffink’s objection, that the conventional theory of quantum mechanics, as formulated by von Neumann, has no room for the concept of a simultaneous measurement of non-commuting quantities.

However, the most important feature of our approach is the way in which we analyse the distinction between the retrodictive and predictive aspects of a joint measurement process, and the way in which we relate these two aspects to the disturbance produced by the measurement. The distinction between retrodiction and prediction has been discussed by several other authors, as has the reciprocity between the accuracy of a joint measurement, and the disturbance produced [9, 15, 22–24, 26, 32, 33]. However, it appears to us that our method of analysis is a significant source of additional insight. It also has certain technical advantages, in that it provides a numerical characterization of the various contrasting features of a joint measurement process which is particularly concise, and convenient.

In the following sections we will illustrate these points, with an analysis of the Arthurs–Kelly process. This process has been the subject of several previous investigations [2–4, 8, 9, 11, 23]. We derive a number of new results. However, in some cases we only give an alternative derivation of a result which is already known. We thought that this was worth doing since our primary purpose is to illustrate the use of some general methods, which have an application to many other measurement processes.

In section 2 we begin by giving a brief summary of the relevant results from [12]. In section 3, we show how the initial apparatus state may be conveniently expanded in terms of eigenstates of the retrodictive and predictive error operators. We then use this fact to show how the distribution of measured values depends on the distribution of retrodictive errors; and how the final state of the system depends on the distribution of predictive errors. In particular, by appropriately choosing the distribution of retrodictive errors, it is possible to ensure that the distribution of measured values is given by any member of the class of ‘operational distributions’ which are obtained by taking the convolution of two Wigner functions, as discussed by Davies [16], Wódkiewicz [34] and others [9, 28, 35, 36]. With the appropriate choice for the distribution of predictive errors it is possible to prepare the system in any desired pure state.

In section 4 we consider retrodictively optimal processes: i.e. processes which minimize the product of retrodictive errors. We show that in every such case the distribution of results is given by the initial system state Husimi function [37, 38] (or Q representation). This is the result originally obtained by Ali and Prugovečki [18] and, in the context of the Arthurs–Kelly process, by Busch [23].

In section 5 we consider predictively optimal processes, which minimize the product of predictive errors. We show that the distribution of measured values is then related to the final system state anti-Husimi function [38, 39] (or P representation).

The distributions of retrodictive and predictive errors are independent of one another:

it is possible for a measurement to be optimal retrodictively, whilst being very poor predictively, or vice versa. In section 6 we consider completely optimal processes: i.e. processes which are both retrodictively and predictively optimal, and which also minimize the degree of disturbance.

Finally, in section 7, we consider the disturbance of the system by the measurement. It is possible to reduce the level of disturbance below that produced by an optimal measurement, provided one is willing to accept a reduced degree of accuracy. In section 7 we consider measurements which give the maximum degree of accuracy consistent with a given level of disturbance.

Before concluding these introductory remarks we should, perhaps, observe that the contrast we drew above, between the ‘POVM approach’, and the approach adopted in this paper, is potentially misleading. It is, in fact, almost impossible to comment on simultaneous measurement processes without tacitly introducing the concept of a POVM. When we distinguish our own approach from the POVM approach we do not mean to suggest that POVMs are not implicitly present in our analysis. We only mean that the concept does not explicitly play the same central role that it does in, for example, the book by Busch *et al* [9].

In the following we make no direct use of results obtained within the context of the POVM approach. That is because the work reported here is a continuation of the work reported in [12]. The original motive for this work was our desire to answer the first of Uffink’s objections (as itemized above). In order to do so we considered it necessary to go back to first principles, and to think things through from the foundation up. We were thereby led to a different approach to the theory of simultaneous measurements of position and momentum, which it seems most natural to present in the manner in which we discovered it, independently of the POVM approach.

It would obviously be desirable to relate the approach taken in this paper to the POVM approach. However, that is a task which we prefer to leave to a future publication, when we also hope to address the second of Uffink’s objections (as itemized above).

2. Definition of the errors and disturbances

In the process described by Arthurs and Kelly [2] a system, with position \hat{x} and momentum \hat{p} , interacts with an apparatus characterized by two pointer observables $\hat{\mu}_X$ (measuring the value of \hat{x}) and $\hat{\mu}_P$ (measuring the value of \hat{p}). Let $\hat{\pi}_X$ and $\hat{\pi}_P$ be the momenta conjugate to $\hat{\mu}_X$ and $\hat{\mu}_P$ respectively. Then

$$[\hat{x}, \hat{p}] = [\hat{\mu}_X, \hat{\pi}_X] = [\hat{\mu}_P, \hat{\pi}_P] = i\hbar$$

all other commutators between the operators \hat{x} , \hat{p} , $\hat{\mu}_X$, $\hat{\pi}_P$, $\hat{\mu}_P$, $\hat{\pi}_X$ being zero.

The measurement interaction is described by the unitary evolution operator

$$\hat{U} = \exp \left[-\frac{i}{\hbar} (\hat{\pi}_P \hat{p} + \hat{\pi}_X \hat{x}) \right]. \quad (3)$$

We assume that system+apparatus are initially in the product state $|\psi \otimes \phi_{\text{ap}}\rangle$, $|\psi\rangle$ being the initial state of the system, and $|\phi_{\text{ap}}\rangle$ the initial state of the apparatus.

In order to define the errors and disturbances we switch to the Heisenberg picture. Let \hat{O} be any of the Schrödinger picture operators \hat{x} , \hat{p} , $\hat{\mu}_X$, $\hat{\pi}_P$, $\hat{\mu}_P$, $\hat{\pi}_X$. Let $\hat{O}_i = \hat{O}$ be the value of the corresponding Heisenberg picture operator immediately before the interaction.

Let $\hat{\mathcal{O}}_f = \hat{U}^\dagger \hat{\mathcal{O}} \hat{U}$ be its value immediately afterwards. Then

$$\begin{aligned}\hat{x}_f &= \hat{x} + \hat{\pi}_p & \hat{p}_f &= \hat{p} - \hat{\pi}_x \\ \hat{\mu}_{Xf} &= \hat{\mu}_X + \hat{x} + \frac{1}{2}\hat{\pi}_p & \hat{\pi}_{Xf} &= \hat{\pi}_X \\ \hat{\mu}_{Pf} &= \hat{\mu}_P + \hat{p} - \frac{1}{2}\hat{\pi}_X & \hat{\pi}_{Pf} &= \hat{\pi}_X.\end{aligned}\quad (4)$$

Define the operators $\hat{\epsilon}_{Xi}$, $\hat{\epsilon}_{Pi}$, $\hat{\epsilon}_{Xf}$, $\hat{\epsilon}_{Pf}$, $\hat{\delta}_X$, $\hat{\delta}_P$ by

$$\begin{aligned}\hat{\epsilon}_{Xi} &= \hat{\mu}_{Xf} - \hat{x}_i & \hat{\epsilon}_{Pi} &= \hat{\mu}_{Pf} - \hat{p}_i \\ \hat{\epsilon}_{Xf} &= \hat{\mu}_{Xf} - \hat{x}_f & \hat{\epsilon}_{Pf} &= \hat{\mu}_{Pf} - \hat{p}_f \\ \hat{\delta}_X &= \hat{x}_f - \hat{x}_i & \hat{\delta}_P &= \hat{p}_f - \hat{p}_i.\end{aligned}\quad (5)$$

Following Braginsky and Khalili [31] we then define the rms errors of retrodiction

$$\Delta_{ei}x = (\langle \psi \otimes \phi_{ap} | \hat{\epsilon}_{Xi}^2 | \psi \otimes \phi_{ap} \rangle)^{\frac{1}{2}} \quad \Delta_{ei}p = (\langle \psi \otimes \phi_{ap} | \hat{\epsilon}_{Pi}^2 | \psi \otimes \phi_{ap} \rangle)^{\frac{1}{2}} \quad (6)$$

and the rms disturbances

$$\Delta_d x = (\langle \psi \otimes \phi_{ap} | \hat{\delta}_X^2 | \psi \otimes \phi_{ap} \rangle)^{\frac{1}{2}} \quad \Delta_d p = (\langle \psi \otimes \phi_{ap} | \hat{\delta}_P^2 | \psi \otimes \phi_{ap} \rangle)^{\frac{1}{2}}.$$

$\Delta_{ei}x$, $\Delta_{ei}p$ correspond to the quantities $\Delta x_{\text{measure}}$, $\Delta p_{\text{measure}}$ defined by Braginsky and Khalili in the context of single measurements of \hat{x} or \hat{p} only. They provide a numerical indication of the accuracy with which the result of the measurement reflects the initial state of the system. $\Delta_d x$, $\Delta_d p$ correspond to Braginsky and Khalili's $\Delta x_{\text{perturb}}$, $\Delta p_{\text{perturb}}$. They provide a numerical indication of the degree in which the measurement perturbs the state of the system.

We also define the rms errors of prediction

$$\Delta_{ef}x = (\langle \psi \otimes \phi_{ap} | \hat{\epsilon}_{Xf}^2 | \psi \otimes \phi_{ap} \rangle)^{\frac{1}{2}} \quad \Delta_{ef}p = (\langle \psi \otimes \phi_{ap} | \hat{\epsilon}_{Pf}^2 | \psi \otimes \phi_{ap} \rangle)^{\frac{1}{2}}.$$

These quantities provide a numerical indication of the accuracy with which the result of the measurement reflects the final state of the system. Braginsky and Khalili do not consider this second kind of error.

We have given a detailed discussion of the interpretation of the quantities $\Delta_{ei}x$, $\Delta_{ei}p$, $\Delta_{ef}x$, $\Delta_{ef}p$, $\Delta_d x$, $\Delta_d p$ in [12].

In view of equations (4) and (5) we have

$$\begin{aligned}\hat{\epsilon}_{Xi} &= \hat{\mu}_X + \frac{1}{2}\hat{\pi}_p & \hat{\epsilon}_{Pi} &= \hat{\mu}_P - \frac{1}{2}\hat{\pi}_X \\ \hat{\epsilon}_{Xf} &= \hat{\mu}_X - \frac{1}{2}\hat{\pi}_p & \hat{\epsilon}_{Pf} &= \hat{\mu}_P + \frac{1}{2}\hat{\pi}_X \\ \hat{\delta}_X &= \hat{\pi}_p & \hat{\delta}_P &= -\hat{\pi}_X.\end{aligned}\quad (7)$$

Consequently

$$\begin{aligned}[\hat{\epsilon}_{Xi}, \hat{\epsilon}_{Pi}] &= -i\hbar & [\hat{\epsilon}_{Xf}, \hat{\epsilon}_{Pf}] &= i\hbar \\ [\hat{\epsilon}_{Xi}, \hat{\delta}_P] &= -i\hbar & [\hat{\epsilon}_{Xf}, \hat{\delta}_P] &= -i\hbar \\ [\hat{\delta}_X, \hat{\epsilon}_{Pi}] &= -i\hbar & [\hat{\delta}_X, \hat{\epsilon}_{Pf}] &= -i\hbar\end{aligned}$$

all other commutators between the operators $\hat{\epsilon}_{Xi}$, $\hat{\epsilon}_{Pi}$, $\hat{\epsilon}_{Xf}$, $\hat{\epsilon}_{Pf}$, $\hat{\delta}_X$, $\hat{\delta}_P$ being zero. We deduce the retrodictive and predictive error relationships

$$\Delta_{ei}x \Delta_{ei}p \geq \frac{\hbar}{2} \quad (8)$$

$$\Delta_{ef}x \Delta_{ef}p \geq \frac{\hbar}{2} \quad (9)$$

and the four error-disturbance relationships

$$\begin{aligned} \Delta_{ei}x \Delta_d p &\geq \frac{\hbar}{2} & \Delta_{ef}x \Delta_d p &\geq \frac{\hbar}{2} \\ \Delta_{ei}p \Delta_d x &\geq \frac{\hbar}{2} & \Delta_{ef}p \Delta_d x &\geq \frac{\hbar}{2}. \end{aligned} \quad (10)$$

Equations (8) and (9) jointly comprise a precise, quantitative statement of the well known principle, that the product of the errors in a simultaneous measurement of position and momentum must be greater than a number $\sim \hbar$. This principle is logically distinct from the uncertainty principle usually so-called [13, 14, 33, 40].

In [12] we have shown that equations (8)–(10) hold for many other simultaneous measurement processes, apart from the Arthurs–Kelly process.

The distinction between the two different aspects of a quantum mechanical measurement process—the retrodictive or determinative aspect *versus* the predictive or preparative one—has been discussed by numerous authors, as has the unavoidable perturbation of the system by the measurement [9, 15, 23, 24, 26, 32, 33]. The quantities $\Delta_{ei}x$, $\Delta_{ei}p$, $\Delta_{ef}x$, $\Delta_{ef}p$, $\Delta_d x$, $\Delta_d p$, and the inequalities relating them, provide a convenient numerical characterization of these features. However, it ought to be stressed that the characterization is not complete. It is, for instance, clearly impossible to give an exhaustive description of the change in the state of the system by only specifying two numbers. The values of the six quantities defined in this section encapsulate some important properties of the measurement process. It is not to be supposed that they encapsulate every significant property.

The method of describing statistical quantities in terms of variances, and mean square values—the method adopted here, in other words—is subject to certain limitations. Recently there has accordingly been some interest in devising alternative approaches. One approach is that involving parameter-based uncertainty relationships [33, 41]. Another approach is that involving entropic uncertainty relationships [24–26, 42]. It would be interesting to see if either or both of these approaches could be used to develop the formalism of this paper.

It will be observed, that the retrodictive and predictive rms errors enter equations (8)–(10) in a completely symmetric manner. As we shall see the two kinds of error can also be fixed independently of one another (in the case of the Arthurs–Kelly process). With the appropriate choice of initial apparatus state it is possible to arrange for the retrodictive errors to be small while the predictive errors are large; or for the predictive errors to be small while the retrodictive errors are large; or for the retrodictive and predictive errors both to be small. In particular, there exists an initial apparatus state for which $\Delta_{ei}x = \Delta_{ef}x$, $\Delta_{ei}p = \Delta_{ef}p$ and for which the products $\Delta_{ei}x \Delta_{ei}p$ and $\Delta_{ef}x \Delta_{ef}p$ both take the minimum value of $\frac{\hbar}{2}$. We shall refer to such measurements as completely optimal. They are discussed in section 6.

It should, however, be stressed that this symmetry between the retrodictive and predictive aspects of a simultaneous measurement process disappears when one turns to the physical interpretation of the quantities $\Delta_{ei}x$, $\Delta_{ei}p$, $\Delta_{ef}x$, $\Delta_{ef}p$, as discussed in section 5 of [12]. One should also bear in mind the fact just mentioned, that these quantities only yield a partial characterization of the measurement process. When we refer to them as the rms errors of retrodiction and prediction we mean, strictly and precisely, no more, and no less than what is stated in section 5 of [12].

The subject of retrodiction and prediction, and the relationship between them, raises some deep conceptual questions which have been the subject of discussion ever since the formulation of the modern theory of quantum mechanics in the mid-1920s [43]. It appears to us that the methods developed in [12], and in this paper, might be used to gain some

additional insight into these questions. However, there is clearly a great deal more work which needs to be done.

3. The distribution of measured values

We see from equation (7) that the error and disturbance operators only depend on the apparatus observables $\hat{\mu}_X$, $\hat{\pi}_X$, $\hat{\mu}_P$, $\hat{\pi}_P$. It follows that the rms errors and disturbances are independent of the initial system state $|\psi\rangle$. We also see that the operators $\hat{\epsilon}_{Xi}$, $\hat{\epsilon}_{Xf}$ constitute a complete commuting set for the apparatus state space, with conjugate momenta $-\hat{\epsilon}_{Pi}$, $\hat{\epsilon}_{Pf}$. It will be convenient to work in terms of simultaneous eigenkets of $\hat{\epsilon}_{Xi}$ and $\hat{\epsilon}_{Xf}$, which we denote $|\epsilon_{Xi}, \epsilon_{Xf}\rangle_{\hat{\epsilon}_{Xi}, \hat{\epsilon}_{Xf}}$. They are related to the simultaneous eigenkets of $\hat{\mu}_X$ and $\hat{\pi}_P$, denoted $|\mu_X, \pi_P\rangle_{\hat{\mu}_X, \hat{\pi}_P}$, by

$$|\epsilon_{Xi}, \epsilon_{Xf}\rangle_{\hat{\epsilon}_{Xi}, \hat{\epsilon}_{Xf}} = \frac{1}{2}(\epsilon_{Xi} + \epsilon_{Xf}), (\epsilon_{Xi} - \epsilon_{Xf})_{\hat{\mu}_X, \hat{\pi}_P}.$$

We now use this equation, equation (3) and the Baker–Campbell–Hausdorff identity to deduce

$$\begin{aligned} \hat{x}, \hat{\mu}_X, \hat{\pi}_P \langle x, \mu_X, \pi_P | \hat{U} | \psi \otimes \phi_{ap} \rangle \\ = \hat{x}, \hat{\mu}_X, \hat{\pi}_P \langle x, \mu_X, \pi_P | \exp\left(-\frac{i}{2\hbar} \hat{\pi}_P \hat{\pi}_X\right) \exp\left(-\frac{i}{\hbar} \hat{\pi}_P \hat{p}\right) \exp\left(-\frac{i}{\hbar} \hat{\pi}_X \hat{x}\right) | \psi \otimes \phi_{ap} \rangle \\ = \hat{x} \langle x - \pi_P | \psi \rangle_{\hat{\epsilon}_{Xi}, \hat{\epsilon}_{Xf}} (\mu_X + \pi_P - x), (\mu_X - x) | \phi_{ap} \rangle. \end{aligned}$$

Taking Fourier transforms we obtain

$$\begin{aligned} \hat{x}, \hat{\mu}_X, \hat{\mu}_P \langle x, \mu_X, \mu_P | \hat{U} | \psi \otimes \phi_{ap} \rangle \\ = \sqrt{\frac{1}{\hbar}} \int dx' \exp\left[\frac{i}{\hbar} \mu_P (x - x')\right]_{\hat{\epsilon}_{Xi}, \hat{\epsilon}_{Xf}} (\mu_X - x'), (\mu_X - x) | \phi_{ap} \rangle_{\hat{x}} \langle x' | \psi \rangle. \quad (11) \end{aligned}$$

We can now calculate $\rho(\mu_X, \mu_P)$, the probability density function describing the result of the measurement:

$$\begin{aligned} \rho(\mu_X, \mu_P) &= \int dx |\hat{x}, \hat{\mu}_X, \hat{\mu}_P \langle x, \mu_X, \mu_P | \hat{U} | \psi \otimes \phi_{ap} \rangle|^2 \\ &= \frac{1}{\hbar} \int dx' dx'' \exp\left[\frac{i}{\hbar} \mu_P (x'' - x')\right] \langle \mu_X - x' | \hat{\rho}_{\epsilon_i} | \mu_X - x'' \rangle \langle x' | \psi \rangle \langle \psi | x'' \rangle \end{aligned} \quad (12)$$

where $\hat{\rho}_{\epsilon_i}$ is the reduced initial apparatus state density matrix corresponding to the pair $\hat{\epsilon}_{Xi}$, $-\hat{\epsilon}_{Pi}$:

$$\langle \epsilon_{Xi} | \hat{\rho}_{\epsilon_i} | \epsilon'_{Xi} \rangle = \int d\epsilon_{Xf} \langle \epsilon_{Xi}, \epsilon_{Xf} | \phi_{ap} \rangle \langle \phi_{ap} | \epsilon'_{Xi}, \epsilon_{Xf} \rangle_{\hat{\epsilon}_{Xi}, \hat{\epsilon}_{Xf}}.$$

Let $W_{sy,i}$ be the Wigner function describing the initial state of the system, and let W_{ϵ_i} be the Wigner function corresponding to $\hat{\rho}_{\epsilon_i}$:

$$\begin{aligned} W_{sy,i}(x, p) &= \frac{1}{\hbar} \int dy \exp\left(\frac{i}{\hbar} py\right) \langle x - \frac{1}{2}y | \psi \rangle \langle \psi | x + \frac{1}{2}y \rangle \\ W_{\epsilon_i}(\epsilon_{Xi}, \epsilon_{Pi}) &= \frac{1}{\hbar} \int dy \exp\left(-\frac{i}{\hbar} \epsilon_{Pi} y\right) \langle \epsilon_{Xi} - \frac{1}{2}y | \hat{\rho}_{\epsilon_i} | \epsilon_{Xi} + \frac{1}{2}y \rangle. \end{aligned}$$

Then the distribution of measured values can be written

$$\rho(\mu_X, \mu_P) = \int dx dp W_{\epsilon_i}(\mu_X - x, \mu_P - p) W_{\text{sy},i}(x, p). \quad (13)$$

With a suitable choice for the distribution of retrodictive errors, it is possible to obtain any member of the class of operational phase-space distributions, discussed by Davies [16], Wódkiewicz [34] and others [9, 28, 35, 36].

The fact, that the Arthurs–Kelly process can be used to obtain any member of this class of distributions is shown in Leonhardt [11] (see also Ban [28]). The novelty of the derivation just given consists in the fact that we have shown that the distribution of measured values depends only on the distribution of retrodictive errors, and is independent of the distribution of predictive errors.

In certain cases the convolution in equation (13) can be inverted [24, 35, 44]. This means, that the original state can be reconstructed from the measured probability distribution *provided* that the latter is known with perfect accuracy—a fact which is sometimes expressed by saying that the measurement is informationally complete [9, 24]. However, it should be observed that the fact is of less practical usefulness than may initially appear owing to the amplification of statistical errors which occurs when one actually tries to carry out the inversion using real experimental data [45].

The right-hand side of equation (13) only depends on the distribution of retrodictive errors. If, on the other hand, one wants to relate $\rho(\mu_X, \mu_P)$ to the final state of the system, then one needs to consider the distribution of predictive errors. We confine ourselves to the case when the initial apparatus state factorizes:

$$\langle \epsilon_{X_i}, \epsilon_{X_f} | \phi_{\text{ap}} \rangle = \langle \epsilon_{X_i} | \phi_{\epsilon_i} \rangle \langle \epsilon_{X_f} | \phi_{\epsilon_f} \rangle. \quad (14)$$

Suppose that the pointer positions μ_X, μ_P are recorded to be in a region \mathcal{R} . Let $\hat{\rho}_{\text{sy},f}$ be the reduced density matrix representing the state of the system immediately after the measurement. Then

$$\langle x | \hat{\rho}_{\text{sy},f} | x' \rangle = \frac{1}{p_{\mathcal{R}}} \int_{\mathcal{R}} d\mu_X d\mu_P \langle x, \mu_X, \mu_P | \hat{U} | \psi \otimes \phi_{\text{ap}} \rangle \langle \psi \otimes \phi_{\text{ap}} | \hat{U} | x', \mu_X, \mu_P \rangle$$

where $p_{\mathcal{R}}$ is the probability of finding μ_X and μ_P in the region \mathcal{R} :

$$p_{\mathcal{R}} = \int_{\mathcal{R}} d\mu_X d\mu_P \rho(\mu_X, \mu_P).$$

Using equations (11), (12) and (14) we obtain

$$\begin{aligned} \langle x | \hat{\rho}_{\text{sy},f} | x' \rangle &= \frac{1}{p_{\mathcal{R}}} \int_{\mathcal{R}} d\mu_X d\mu_P \exp \left[\frac{i}{\hbar} \mu_P (x - x') \right] \\ &\times \rho(\mu_X, \mu_P) \hat{\epsilon}_{X_f} \langle \mu_X - x | \phi_{\epsilon_f} \rangle \langle \phi_{\epsilon_f} | \mu_X - x' \rangle \hat{\epsilon}_{X_f}. \end{aligned} \quad (15)$$

Let $W_{\text{sy},f}$ be the Wigner function describing the final state of the system, and let W_{ϵ_f} be the Wigner function corresponding to the state $|\phi_{\epsilon_f}\rangle$:

$$\begin{aligned} W_{\text{sy},f}(x, p) &= \frac{1}{h} \int dy \exp \left(\frac{i}{\hbar} p y \right) \langle x - \frac{1}{2} y | \hat{\rho}_{\text{sy},f} | x + \frac{1}{2} y \rangle \\ W_{\epsilon_f}(\epsilon_{X_f}, \epsilon_{P_f}) &= \frac{1}{h} \int dy \exp \left(\frac{i}{\hbar} \epsilon_{P_f} y \right) \langle \epsilon_{X_f} - \frac{1}{2} y | \phi_{\epsilon_f} \rangle \langle \phi_{\epsilon_f} | \epsilon_{X_f} + \frac{1}{2} y \rangle. \end{aligned}$$

Then equation (15) becomes

$$W_{\text{sy},f}(x, p) = \frac{1}{p_{\mathcal{R}}} \int_{\mathcal{R}} d\mu_X d\mu_P W_{\epsilon_f}(\mu_X - x, \mu_P - p) \rho(\mu_X, \mu_P). \quad (16)$$

Equation (13) shows how the distribution of retrodictive errors can be used to express ρ in terms of $W_{\text{sy},i}$. Equation (16) shows how the distribution of predictive errors can be used to express $W_{\text{sy},f}$ in terms of ρ .

If \mathcal{R} is a sufficiently small region surrounding the point (μ_X, μ_P)

$$W_{\text{sy},f}(x, p) \approx W_{\epsilon_f}(\mu_X - x, \mu_P - p).$$

We see, that with a suitable choice for the distribution of predictive errors, the Arthurs–Kelly process can be used to prepare the system in any desired pure state.

4. Retrodictively optimal measurements

Suppose that the measurement maximizes the degree of retrodictive accuracy:

$$\Delta_{\text{ei}x} \Delta_{\text{ei}p} = \frac{\hbar}{2}. \quad (17)$$

Define the quantity λ_i by

$$\Delta_{\text{ei}x} = \frac{\lambda_i}{\sqrt{2}} \quad \Delta_{\text{ei}p} = \frac{\hbar}{\sqrt{2}\lambda_i}.$$

We shall refer to λ_i as the retrodictive spatial resolution.

The necessary and sufficient condition for equation (17) to be true is that the initial apparatus wavefunction be of the form

$$\langle \epsilon_{X_i}, \epsilon_{X_f} | \phi_{\text{ap}} \rangle = \left(\frac{1}{\pi \lambda_i^2} \right)^{\frac{1}{4}} \exp \left(-\frac{1}{2\lambda_i^2} \epsilon_{X_i}^2 \right) \phi_f(\epsilon_{X_f}) \quad (18)$$

where $\phi_f(\epsilon_{X_f})$ is an arbitrary normalized function. ϕ_f determines the errors of prediction. The fact that it is arbitrary means, that requiring the measurement to be retrodictively optimal places no constraint on the predictive accuracy. The two kinds of error are completely independent.

In the $\hat{\pi}_X, \hat{\mu}_P$ -representation equation (18) takes the form

$$\langle \pi_X, \mu_P | \phi_{\text{ap}} \rangle = \left(\frac{\lambda_i^2}{\pi \hbar^2} \right)^{\frac{1}{4}} \exp \left[-\frac{\lambda_i^2}{2\hbar^2} (\mu_P - \frac{1}{2}\pi_X)^2 \right] \tilde{\phi}_f(\mu_P + \frac{1}{2}\pi_X)$$

where $\tilde{\phi}_f$ is the Fourier transform of ϕ_f :

$$\tilde{\phi}_f(\epsilon_{P_f}) = \sqrt{\frac{1}{\hbar}} \int d\epsilon_{X_f} \exp \left(-\frac{i}{\hbar} \epsilon_{P_f} \epsilon_{X_f} \right) \phi_f(\epsilon_{X_f}).$$

We recognize the wavefunction considered by Stenholm [4].

Using equation (13) we find, for the probability distribution of measured values,

$$\rho(\mu_X, \mu_P) = Q_{i,\lambda_i}(\mu_X, \mu_P)$$

where Q_{i,λ_i} is the initial system state Husimi function [37, 38]:

$$Q_{i,\lambda_i}(\mu_X, \mu_P) = \frac{2}{\hbar} \int dx dp \exp \left[-\frac{1}{\lambda_i^2} (\mu_X - x)^2 - \frac{\lambda_i^2}{\hbar^2} (\mu_P - p)^2 \right] W_{\text{sy},i}(x, p).$$

The fact that the Husimi function gives the distribution of measured values for the case of maximal accuracy was first shown by Ali and Prugovečki [18], working in terms of the approach based on POVMs and unsharp observables. The specialization of this result to the case of the Arthurs–Kelly process was first discussed by Busch [23].

Finally, let us calculate the disturbances in this case. We have

$$\begin{aligned} (\Delta_{dX})^2 &= \int d\epsilon_{Xi} d\epsilon_{Xf} (\epsilon_{Xi} - \epsilon_{Xf})^2 |\langle \epsilon_{Xi}, \epsilon_{Xf} | \phi_{ap} \rangle|^2 = \frac{\lambda_i^2}{2} + (\Delta_{efX})^2 \\ (\Delta_{dP})^2 &= \int d\epsilon_{Pi} d\epsilon_{Pf} (\epsilon_{Pi} - \epsilon_{Pf})^2 |\langle \epsilon_{Pi}, \epsilon_{Pf} | \phi_{ap} \rangle|^2 = \frac{\hbar^2}{2\lambda_i^2} + (\Delta_{efP})^2. \end{aligned} \tag{19}$$

Using the predictive error relationship, equation (9), we deduce

$$\Delta_{dX} \Delta_{dP} \geq \hbar.$$

5. Predictively optimal measurements

We next consider measurements which maximize the predictive accuracy:

$$\Delta_{efX} \Delta_{efP} = \frac{\hbar}{2}. \tag{20}$$

Define the quantity λ_f by

$$\Delta_{efX} = \frac{\lambda_f}{\sqrt{2}} \quad \Delta_{efP} = \frac{\hbar}{\sqrt{2}\lambda_f}.$$

We shall refer to λ_f as the predictive spatial resolution. The necessary and sufficient condition for equation (20) to be true is that $\langle \epsilon_{Xi}, \epsilon_{Xf} | \phi_{ap} \rangle$ be of the form

$$\langle \epsilon_{Xi}, \epsilon_{Xf} | \phi_{ap} \rangle = \left(\frac{1}{\pi \lambda_f^2} \right)^{\frac{1}{4}} \exp \left(-\frac{1}{2\lambda_f^2} \epsilon_{Xf}^2 \right) \phi_i(\epsilon_{Xi}). \tag{21}$$

Suppose that the final pointer positions are recorded to be in the region \mathcal{R} . In view of equation (15) the final system state reduced density matrix is given by

$$\hat{\rho}_{sy,f} = \frac{1}{P_{\mathcal{R}}} \int_{\mathcal{R}} d\mu_X d\mu_P \rho(\mu_X, \mu_P) |(\mu_X, \mu_P)_{\lambda_f}\rangle \langle (\mu_X, \mu_P)_{\lambda_f}| \tag{22}$$

where $|(\mu_X, \mu_P)_{\lambda_f}\rangle$ is the coherent state with wavefunction

$$\langle x | (\mu_X, \mu_P)_{\lambda_f} \rangle = \left(\frac{1}{\pi \lambda_f^2} \right)^{\frac{1}{4}} \exp \left[-\frac{1}{2\lambda_f^2} (x - \mu_X)^2 + \frac{i}{\hbar} \mu_P x - \frac{i}{2\hbar} \mu_P \mu_X \right].$$

Let P_{f,λ_f} be the anti-Husimi function describing the final state of the system (the P representation of Glauber and Sudarshan). We have [38, 39]

$$\hat{\rho}_{sy,f} = \int d\mu_X d\mu_P P_{f,\lambda_f}(\mu_X, \mu_P) |(\mu_X, \mu_P)_{\lambda_f}\rangle \langle (\mu_X, \mu_P)_{\lambda_f}|. \tag{23}$$

Comparing equations (22) and (23) we deduce

$$P_{f,\lambda_f}(\mu_X, \mu_P) = \begin{cases} \frac{1}{P_{\mathcal{R}}} \rho(\mu_X, \mu_P) & \text{if } (\mu_X, \mu_P) \in \mathcal{R} \\ 0 & \text{otherwise.} \end{cases}$$

If \mathcal{R} is a sufficiently small region surrounding the point (μ_X, μ_P) , then the system is approximately in the state $|(\mu_X, \mu_P)_{\lambda_f}\rangle$ after the measurement:

$$\hat{\rho}_{\text{sy},f} \approx |(\mu_X, \mu_P)_{\lambda_f}\rangle\langle(\mu_X, \mu_P)_{\lambda_f}|.$$

The reader may easily verify that

$$\Delta_{\text{d}x}\Delta_{\text{d}p} \geq \hbar$$

as in the case of a retrodictively optimal process.

6. Completely optimal measurements

Suppose that the measurement is both retrodictively optimal at spatial resolution λ_i , and predictively optimal at spatial resolution λ_f . In view of equations (18) and (21) the initial apparatus wavefunction must be

$$\langle\epsilon_{X_i}, \epsilon_{X_f}|\phi_{\text{ap}}\rangle = (\pi\lambda_i\lambda_f)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\lambda_i^2}\epsilon_{X_i}^2 - \frac{1}{2\lambda_f^2}\epsilon_{X_f}^2\right).$$

We have from equation (19)

$$\Delta_{\text{d}x}\Delta_{\text{d}p} = \frac{\hbar}{2} \left(2 + \frac{\lambda_f^2}{\lambda_i^2} + \frac{\lambda_i^2}{\lambda_f^2}\right)^{\frac{1}{2}} \geq \hbar.$$

The necessary and sufficient condition for this expression to achieve its lower bound is that both the retrodiction and prediction should be at the same spatial resolution: $\lambda_i = \lambda_f = \lambda$, say. We then have, in the $\hat{\mu}_X, \hat{\mu}_P$ representation

$$\langle\mu_X, \mu_P|\phi_{\text{ap}}\rangle = \frac{2}{\sqrt{\hbar}} \exp\left(-\frac{1}{\lambda^2}\mu_X^2 - \frac{\lambda^2}{\hbar^2}\mu_P^2\right)$$

which is the wavefunction considered by Arthurs and Kelly [2]. With this choice of $|\phi_{\text{ap}}\rangle$ the process produces the least amount of disturbance consistent with maximal accuracy. It might therefore be described as a completely optimal process.

It is interesting to note, however, that λ_i and λ_f are completely independent. One could, for instance, have $\lambda_i \rightarrow 0$ and $\lambda_f \rightarrow \infty$ —so that the measurement gives an almost perfectly accurate retrodiction of position, whilst at the same time preparing the system in an approximate eigenstate of momentum. Of course, the disturbances would then be very large.

7. Minimally disturbing measurements

It is possible to make $\Delta_{\text{d}x}\Delta_{\text{d}p}$ smaller than \hbar , provided that one is willing to accept some loss of accuracy. In this section we address the question: what is the greatest accuracy which can be achieved for a given level of disturbance? We confine ourselves to the case when the retrodictive and predictive errors are equal:

$$\Delta_{\text{ei}x} = \Delta_{\text{ef}x} \quad \text{and} \quad \Delta_{\text{ei}p} = \Delta_{\text{ef}p}. \quad (24)$$

We assume that the product of disturbances is given by

$$\Delta_{\text{d}x}\Delta_{\text{d}p} = \hbar e^{-\eta} \quad (25)$$

for some $\eta \geq 0$. We then ask: what is the least value of the product $\Delta_{ei}x \Delta_{ei}p = \Delta_{ef}x \Delta_{ef}p$ subject to this constraint? And what is the probability distribution of measured values when the lower bound is achieved?

It is convenient to define

$$\hat{\epsilon}_X = \frac{1}{2}(\hat{\epsilon}_{Xi} + \hat{\epsilon}_{Xf}) \quad \text{and} \quad \hat{\epsilon}_P = \frac{1}{2}(\hat{\epsilon}_{Pi} + \hat{\epsilon}_{Pf}). \quad (26)$$

We also have

$$\hat{\delta}_X = \hat{\epsilon}_{Xi} - \hat{\epsilon}_{Xf} \quad \text{and} \quad \hat{\delta}_P = \hat{\epsilon}_{Pi} - \hat{\epsilon}_{Pf}. \quad (27)$$

Consequently

$$[\hat{\epsilon}_X, \hat{\delta}_P] = [\hat{\delta}_X, \hat{\epsilon}_P] = -i\hbar$$

all other commutators between $\hat{\epsilon}_X, \hat{\delta}_X, \hat{\epsilon}_P, \hat{\delta}_P$ being zero. We see that $\hat{\epsilon}_X, \hat{\delta}_X$ constitute a complete commuting set of apparatus observables, with conjugate momenta $-\hat{\delta}_P, -\hat{\epsilon}_P$. In particular

$$\begin{aligned} \langle \phi_{ap} | \hat{\epsilon}_X^2 | \phi_{ap} \rangle \langle \phi_{ap} | \hat{\delta}_P^2 | \phi_{ap} \rangle &\geq \frac{\hbar^2}{4} \\ \langle \phi_{ap} | \hat{\delta}_X^2 | \phi_{ap} \rangle \langle \phi_{ap} | \hat{\epsilon}_P^2 | \phi_{ap} \rangle &\geq \frac{\hbar^2}{4}. \end{aligned} \quad (28)$$

It follows from equations (26) and (27)

$$\begin{aligned} (\Delta_{ei}x)^2 &= \langle \phi_{ap} | \hat{\epsilon}_X^2 | \phi_{ap} \rangle + \frac{1}{4} \langle \phi_{ap} | \hat{\delta}_X^2 | \phi_{ap} \rangle + \langle \phi_{ap} | \hat{\epsilon}_X \hat{\delta}_X | \phi_{ap} \rangle \\ (\Delta_{ef}x)^2 &= \langle \phi_{ap} | \hat{\epsilon}_X^2 | \phi_{ap} \rangle + \frac{1}{4} \langle \phi_{ap} | \hat{\delta}_X^2 | \phi_{ap} \rangle - \langle \phi_{ap} | \hat{\epsilon}_X \hat{\delta}_X | \phi_{ap} \rangle. \end{aligned}$$

In view of equation (24) we must have

$$\langle \phi_{ap} | \hat{\epsilon}_X \hat{\delta}_X | \phi_{ap} \rangle = 0$$

and

$$(\Delta_{ei}x)^2 = (\Delta_{ef}x)^2 = \langle \phi_{ap} | \hat{\epsilon}_X^2 | \phi_{ap} \rangle + \frac{1}{4} \langle \phi_{ap} | \hat{\delta}_X^2 | \phi_{ap} \rangle. \quad (29)$$

Similarly

$$(\Delta_{ei}p)^2 = (\Delta_{ef}p)^2 = \langle \phi_{ap} | \hat{\epsilon}_P^2 | \phi_{ap} \rangle + \frac{1}{4} \langle \phi_{ap} | \hat{\delta}_P^2 | \phi_{ap} \rangle. \quad (30)$$

In view of equation (25) we can choose λ such that

$$\begin{aligned} \Delta_d x &= \lambda \exp\left(-\frac{\eta}{2}\right) \\ \Delta_d p &= \frac{\hbar}{\lambda} \exp\left(-\frac{\eta}{2}\right). \end{aligned}$$

In view of equation (28) we must then have

$$\begin{aligned} \langle \phi_{ap} | \hat{\epsilon}_X^2 | \phi_{ap} \rangle &\geq \frac{\lambda^2}{4} e^\eta \\ \langle \phi_{ap} | \hat{\epsilon}_P^2 | \phi_{ap} \rangle &\geq \frac{\hbar^2}{4\lambda^2} e^\eta. \end{aligned} \quad (31)$$

Inserting these results in equations (29) and (30) gives

$$\begin{aligned}(\Delta_{ei}x)^2 &= (\Delta_{ef}x)^2 \geq \frac{\lambda^2}{2} \cosh \eta \\ (\Delta_{ei}p)^2 &= (\Delta_{ef}p)^2 \geq \frac{\hbar^2}{2\lambda^2} \cosh \eta\end{aligned}$$

whence

$$\Delta_{ei}x \Delta_{ei}p = \Delta_{ef}x \Delta_{ef}p \geq \frac{\hbar}{2} \cosh \eta$$

which is the desired inequality.

The product of errors achieves its lower bound if and only if the lower bounds set by equation (31) are achieved, so that $|\phi_{ap}\rangle$ is a minimum uncertainty state with respect to the pairs $\hat{\epsilon}_X, -\hat{\delta}_P$ and $\hat{\delta}_X, -\hat{\epsilon}_P$. In the $\hat{\epsilon}_X, \hat{\delta}_X$ -representation

$$\langle \epsilon_X, \delta_X | \phi_{ap} \rangle = \frac{1}{\sqrt{\pi\lambda}} \exp \left[-\frac{1}{\lambda^2} (e^{-\eta} \epsilon_X^2 + \frac{1}{4} e^{\eta} \delta_X^2) \right].$$

Transforming to the $\hat{\epsilon}_{Xi}, \hat{\epsilon}_{Xf}$ -representation we find

$$\langle \epsilon_{Xi}, \epsilon_{Xf} | \phi_{ap} \rangle = \frac{1}{\sqrt{\pi\lambda}} \exp \left[-\frac{1}{2\lambda^2} (\cosh \eta \epsilon_{Xi}^2 - 2 \sinh \eta \epsilon_{Xi} \epsilon_{Xf} + \cosh \eta \epsilon_{Xf}^2) \right].$$

Using equation (13) we obtain the probability distribution of measured values

$$\begin{aligned}\rho(\mu_X, \mu_P) &= \frac{2}{h \cosh \eta} \int dx dp \exp \left[-\frac{1}{\cosh \eta} \left(\frac{1}{\lambda^2} (\mu_X - x)^2 + \frac{\lambda^2}{\hbar^2} (\mu_P - p)^2 \right) \right] \\ &\quad \times W_{i, sy}(x, p).\end{aligned}$$

This is a smeared Wigner function, of the kind proposed by Cartwright [35, 36, 46]. For $\eta = 0$ it reduces to the Husimi function. As η increases $W_{sy,i}$ is smoothed over increasingly large regions of phase space—in agreement with the fact, that the larger η , the less accurate the measurement.

8. Conclusion

In this paper we have only considered the Arthurs–Kelly process. It would be interesting to know whether the results obtained generalize to some of the other measurement processes which have been discussed in the literature [5–7, 10, 11]. We hope to return to this question in a future publication.

Appendix. The definition of the retrodictive errors

In equation (6) we defined the rms errors of retrodiction $\Delta_{ei}x, \Delta_{ei}p$ in terms of the retrodictive error operators $\hat{\epsilon}_{Xi}, \hat{\epsilon}_{Pi}$. The purpose of this appendix is to indicate the relationship between these quantities and the definition of the measurement inaccuracies which was given by Ali and Prugovečki [9, 18].

Consider a measurement process in which a system, with position \hat{x} and momentum \hat{p} interacts with an apparatus, characterized by two pointer observables $\hat{\mu}_X, \hat{\mu}_P$. Let ψ and ϕ_{ap} be the initial states of the system and apparatus respectively. Let \hat{U} be the unitary

evolution operator describing the interaction, so that the final state of the system+apparatus is $\hat{U}|\psi \otimes \phi_{\text{ap}}\rangle$ (in the Schrödinger picture). We no longer confine ourselves to the case of the Arthurs–Kelly process, so \hat{U} is not assumed to have the particular form specified by equation (3).

For a given choice of initial apparatus state ϕ_{ap} we have

$$\langle x, \mu_X, \mu_P | \hat{U} | \psi \otimes \phi_{\text{ap}} \rangle = \int dx' K(x, \mu_X, \mu_P; x') \langle x' | \psi \rangle \quad (32)$$

for some kernel K (the choice of K being dependent on the choice of ϕ_{ap}). The unitarity of \hat{U} implies

$$\int dx d\mu_X d\mu_P K(x, \mu_X, \mu_P; x_1) K^*(x, \mu_X, \mu_P; x_2) = \delta(x_1 - x_2). \quad (33)$$

The distribution of measured values is given by

$$\begin{aligned} \rho(\mu_X, \mu_P) &= \int dx |\langle x, \mu_X, \mu_P | \hat{U} | \psi \otimes \phi_{\text{ap}} \rangle|^2 \\ &= \int dx dx'_1 dx'_2 K(x, \mu_X, \mu_P; x'_1) K^*(x, \mu_X, \mu_P; x'_2) \langle x'_1 | \psi \rangle \langle \psi | x'_2 \rangle \end{aligned}$$

or, in terms of the p representation wavefunction,

$$\rho(\mu_X, \mu_P) = \int dp dp'_1 dp'_2 \tilde{K}(p, \mu_X, \mu_P; p'_1) \tilde{K}^*(p, \mu_X, \mu_P; p'_2) \langle p'_1 | \psi \rangle \langle \psi | p'_2 \rangle$$

where

$$\tilde{K}(p, \mu_X, \mu_P; p') = \frac{1}{h^2} \int dx dx' \exp\left[\frac{i}{\hbar}(p'x' - px)\right] K(x, \mu_X, \mu_P; x').$$

The marginal distributions can be written

$$\begin{aligned} \int d\mu_P \rho(\mu_X, \mu_P) &= \int dx'_1 dx'_2 f_X(\mu_X; x'_1, x'_2) \langle x'_1 | \psi \rangle \langle \psi | x'_2 \rangle \\ \int d\mu_X \rho(\mu_X, \mu_P) &= \int dp'_1 dp'_2 f_P(\mu_P; p'_1, p'_2) \langle p'_1 | \psi \rangle \langle \psi | p'_2 \rangle \end{aligned}$$

where

$$\begin{aligned} f_X(\mu_X; x'_1, x'_2) &= \int dx d\mu_P K(x, \mu_X, \mu_P; x'_1) K^*(x, \mu_X, \mu_P; x'_2) \\ f_P(\mu_P; p'_1, p'_2) &= \int dp d\mu_X \tilde{K}(p, \mu_X, \mu_P; p'_1) \tilde{K}^*(p, \mu_X, \mu_P; p'_2). \end{aligned} \quad (34)$$

Ali and Prugovečki [18] confined themselves to the class of processes for which f_X and f_P have the form

$$\begin{aligned} f_X(\mu_X; x'_1, x'_2) &= \chi_X(\mu_X, x'_1) \delta(x'_2 - x'_1) \\ f_P(\mu_P; p'_1, p'_2) &= \chi_P(\mu_P, p'_1) \delta(p'_2 - p'_1) \end{aligned}$$

for some pair of functions χ_X , χ_P , and they focus on the even more restricted class for which

$$\begin{aligned} f_X(\mu_X; x'_1, x'_2) &= \chi_{X0}(\mu_X - x'_1) \delta(x'_2 - x'_1) \\ f_P(\mu_P; p'_1, p'_2) &= \chi_{P0}(\mu_P - p'_1) \delta(p'_2 - p'_1) \end{aligned} \quad (35)$$

for some pair of functions χ_{X0} , χ_{P0} . This assumption is valid in the case of the Arthurs–Kelly process. However, there does not seem to be any reason to expect it to be true generally.

If f_X and f_P do satisfy the condition of equation (35), the marginal probability distributions for μ_X and μ_P can be written as convolutions of the initial system state probability distributions for x and p :

$$\int d\mu_P \rho(\mu_X, \mu_P) = \int dx' \chi_{X0}(\mu_X - x') |\langle x' | \psi \rangle|^2$$

$$\int d\mu_X \rho(\mu_X, \mu_P) = \int dp' \chi_{P0}(\mu_P - p') |\langle p' | \psi \rangle|^2.$$

Ali and Prugovečki [9, 18] defined the measurement inaccuracies in terms of the widths of these convolutions:

$$\sigma_X = \left(\int dx x^2 \chi_{X0}(x) - \left(\int dx x \chi_{X0}(x) \right)^2 \right)^{\frac{1}{2}}$$

$$\sigma_P = \left(\int dp p^2 \chi_{P0}(p) - \left(\int dp p \chi_{P0}(p) \right)^2 \right)^{\frac{1}{2}}. \quad (36)$$

Let us now compare these definitions with the definitions of the rms errors of retrodiction used in this paper. Referring back to equation (5) we see

$$\langle \psi \otimes \phi_{\text{ap}} | \hat{\epsilon}_{X_i}^2 | \psi \otimes \phi_{\text{ap}} \rangle = \langle \psi \otimes \phi_{\text{ap}} | \hat{U}^\dagger \hat{\mu}_X^2 \hat{U} + \hat{U}^\dagger \hat{\mu}_X \hat{U} \hat{x} + \hat{x} \hat{U}^\dagger \hat{\mu}_X \hat{U} + \hat{x}^2 | \psi \otimes \phi_{\text{ap}} \rangle.$$

Using equations (32) and (34) we find

$$\langle \psi \otimes \phi_{\text{ap}} | \hat{U}^\dagger \hat{\mu}_X^2 \hat{U} | \psi \otimes \phi_{\text{ap}} \rangle = \int d\mu_X d\mu_P dx \mu_X^2 |\langle x, \mu_X, \mu_P | \hat{U} | \psi \otimes \phi_{\text{ap}} \rangle|^2$$

$$= \int d\mu_X dx'_1 dx'_2 \mu_X^2 f_X(\mu_X; x'_1, x'_2) \langle x'_1 | \psi \rangle \langle \psi | x'_2 \rangle$$

and

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

$$= \int d\mu_X d\mu_P dx \mu_X \langle x, \mu_X, \mu_P | \hat{U} \hat{x} | \psi \otimes \phi_{\text{ap}} \rangle \langle \psi \otimes \phi_{\text{ap}} | \hat{U}^\dagger | x, \mu_X, \mu_P \rangle$$

$$= \int d\mu_X dx'_1 dx'_2 \mu_X x'_1 f_X(\mu_X; x'_1, x'_2) \langle x'_1 | \psi \rangle \langle \psi | x'_2 \rangle.$$

Similarly

$$\langle \psi \otimes \phi_{\text{ap}} | \hat{x} \hat{U}^\dagger \hat{\mu}_X \hat{U} | \psi \otimes \phi_{\text{ap}} \rangle = \int d\mu_X dx'_1 dx'_2 \mu_X x'_2 f_X(\mu_X; x'_1, x'_2) \langle x'_1 | \psi \rangle \langle \psi | x'_2 \rangle.$$

Finally

$$\langle \psi \otimes \phi_{\text{ap}} | \hat{x}^2 | \psi \otimes \phi_{\text{ap}} \rangle = \int dx' x'^2 |\langle x' | \psi \rangle|^2$$

$$= \int d\mu_X dx'_1 dx'_2 x'_1 x'_2 f_X(\mu_X; x'_1, x'_2) \langle x'_1 | \psi \rangle \langle \psi | x'_2 \rangle$$

where we have used the fact

$$\int d\mu_X f_X(\mu_X; x'_1, x'_2) = \delta(x'_1 - x'_2)$$

as follows from the unitarity condition, equation (33). Putting these results together gives

$$(\Delta_{ei}x)^2 = \int d\mu_X dx'_1 dx'_2 (\mu_X - x'_1)(\mu_X - x'_2) f_X(\mu_X; x'_1, x'_2) \langle x'_1 | \psi \rangle \langle \psi | x'_2 \rangle.$$

In the same way we can derive

$$(\Delta_{ei}p)^2 = \int d\mu_P dp'_1 dp'_2 (\mu_P - p'_1)(\mu_P - p'_2) f_P(\mu_P; p'_1, p'_2) \langle p'_1 | \psi \rangle \langle \psi | p'_2 \rangle.$$

These equations hold quite generally. If the functions f_X , f_P are of the form given by equation (35), then

$$(\Delta_{ei}x)^2 = \int dx x^2 \chi_{X0}(x)$$

$$(\Delta_{ei}p)^2 = \int dp p^2 \chi_{P0}(p).$$

Comparing these equations with equation (36) we see, that the only difference is, that Ali and Prugovečki work in terms of the standard deviations, whereas we use the rms values. Our definition essentially coincides with theirs, therefore, for the class of measurement processes which they consider (and which includes the Arthurs–Kelly process discussed in this paper).

Our definition of the retrodictive errors applies to a larger class of measurement processes than the definition of Ali and Prugovečki. It also has other advantages, since it brings out the connection with the retrodictive error operators. In [12] we used this connection to answer the question of principle raised by Uffink, and to prove the error–error and error–disturbance relationships of equations (8)–(10). In this paper we saw how the expansion in terms of eigenstates of the retrodictive and predictive error operators provides a very convenient way of analysing the mathematical properties of the Arthurs–Kelly process.

References

- [1] Heisenberg W 1927 *Z. Phys.* **43** 172 reprinted in 1983 *Quantum Theory and Measurement* ed J A Wheeler and W H Zurek (Princeton: Princeton University Press)
- [2] Arthurs E and Kelly J L 1965 *Bell Syst. Tech. J.* **44** 725
- [3] Braunstein S L, Caves C M and Milburn G J 1991 *Phys. Rev. A* **43** 1153
- [4] Stenholm S 1992 *Ann. Phys., NY* **218** 233
- [5] Leonhardt U and Paul H 1993 *J. Mod. Opt.* **40** 1745
Leonhardt U and Paul H 1993 *Phys. Rev. A* **48** 4598
- [6] Leonhardt U, Böhmer B and Paul H 1995 *Opt. Commun.* **119** 296
- [7] Törma P, Stenholm S and Jex I 1995 *Phys. Rev. A* **52** 4812
- [8] Power W L, Tan S M and Wilkens M 1997 *J. Mod. Opt.* **44** 2591
- [9] Busch P, Grabowski M and Lahti P J 1995 *Operational Quantum Physics* (Berlin: Springer)
- [10] Leonhardt U and Paul H 1995 *Prog. Quantum Electron.* **19** 89
- [11] Leonhardt U 1997 *Measuring the Quantum State of Light* (Cambridge: Cambridge University Press)
- [12] Appleby D M 1998 *Report no QMW-PH-11, xxx.lanl.gov no quant-ph/9803046 Int. J. Theor. Phys.* to be published
- [13] Wódkiewicz K 1987 *Phys. Lett.* **124A** 207
- [14] Raymer M G 1994 *Am. J. Phys.* **62** 986
- [15] Prugovečki E 1973 *Found. Phys.* **3** 3
Prugovečki E 1976 *J. Math. Phys.* **17** 1673
- [16] Davies E B 1976 *Quantum Theory of Open Systems* (New York: Academic)
- [17] Prugovečki E 1976 *J. Math. Phys.* **17** 517
- [18] Ali S T and Prugovečki E 1977 *J. Math. Phys.* **18** 219

- [19] Prugovečki E 1978 *Ann. Phys., NY* **110** 102
- [20] Holevo A S 1982 *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North-Holland)
- [21] Prugovečki E 1984 *Stochastic Quantum Mechanics and Quantum Space Time* (Dordrecht: Reidel)
- [22] Busch P and Lahti P J 1984 *Phys. Rev. D* **29** 1634
- [23] Busch P 1985 *Int. J. Theor. Phys.* **24** 63
- [24] Busch P and Lahti P J 1989 *Found. Phys.* **19** 633
- [25] Martens H and de Muynck W M 1990 *Found. Phys.* **20** 255
Martens H and de Muynck W M 1990 *Found. Phys.* **20** 357
- [26] Martens H and de Muynck W M 1992 *J. Phys. A: Math. Gen.* **25** 4887
- [27] de Muynck W M, de Baere W and Martens H 1994 *Found. Phys.* **24** 1589
- [28] Ban M 1998 *Int. J. Theor. Phys.* **36** 2583
- [29] Uffink J 1994 *Int. J. Theor. Phys.* **33** 199
- [30] von Neumann J 1932 *Mathematische Grundlagen der Quantenmechanik* (Berlin: Springer) (Engl. transl. *Mathematical Foundations of Quantum Mechanics* (Princeton, NJ: Princeton University Press))
- [31] Braginsky V B and Khalili F Ya 1992 *Quantum Measurement* ed K S Thorne (Cambridge: Cambridge University Press)
- [32] Margenau H 1958 *Phil. Sci.* **25** 23
- [33] Hilgevoord J and Uffink J 1990 *Sixty Two Years of Uncertainty* ed A I Miller (New York: Plenum)
- [34] Wódkiewicz K 1984 *Phys. Rev. Lett.* **52** 1064
- [35] Lalović D, Davidović D M and Bijedić N 1992 *Phys. Rev. A* **46** 1206
- [36] Halliwell J J 1992 *Phys. Rev. D* **46** 1610
- [37] Husimi K 1940 *Proc. Phys. Math. Soc. Japan.* **22** 264
- [38] Hillery M, O'Connell R F, Scully M O and Wigner E P 1984 *Phys. Rep.* **106** 121
Lee H W 1995 *Phys. Rep.* **259** 147
- [39] Glauber R J 1963 *Phys. Rev.* **131** 2766
Sudarshan E C G 1963 *Phys. Rev. Lett.* **10** 277
- [40] Ballentine L E 1970 *Rev. Mod. Phys.* **42** 358
- [41] Braunstein S L, Caves C M and Milburn G J 1996 *Ann. Phys., NY* **247** 135
- [42] Bužek V, Keitel C H and Knight P L 1995 *Phys. Rev. A* **51** 2575
Bužek V, Keitel C H and Knight P L 1995 *Phys. Rev. A* **51** 2594
- [43] Heisenberg W 1930 *The Physical Principles of the Quantum Theory* translated by C Eckart and F C Hoyt (Chicago, IL: University of Chicago Press)
- [44] Wünsche A and Bužek V 1997 *Quantum Semiclass. Opt.* **9** 631
- [45] Leonhardt U and Paul H 1994 *J. Mod. Opt.* **41** 1427
- [46] Cartwright N D 1976 *Physica* **83A** 210
Soto F and Claverie P 1981 *Physica* **109A** 193