

# Heisenberg's Uncertainty Principle

P. Busch\*

*Perimeter Institute for Theoretical Physics, Waterloo, Canada,  
and Department of Mathematics, University of York, York, UK*

T. Heinonen<sup>†</sup> and P.J. Lahti<sup>‡</sup>

*Department of Physics, University of Turku, Turku, Finland*

(Dated: 5 December 2006)

Heisenberg's uncertainty principle is usually taken to express a *limitation* of operational possibilities imposed by quantum mechanics. Here we demonstrate that the full content of this principle also includes its *positive* role as a condition ensuring that mutually exclusive experimental options can be reconciled if an appropriate trade-off is accepted. The uncertainty principle is shown to appear in three manifestations, in the form of uncertainty relations: for the widths of the position and momentum distributions in any quantum state; for the inaccuracies of any joint measurement of these quantities; and for the inaccuracy of a measurement of one of the quantities and the ensuing disturbance in the distribution of the other quantity. Whilst conceptually distinct, these three kinds of uncertainty relations are shown to be closely related formally. Finally, we survey models and experimental implementations of joint measurements of position and momentum and comment briefly on the status of experimental tests of the uncertainty principle.

PACS numbers: 03.65

## Contents

<b>I. Introduction</b>	2
<b>II. From “no joint sharp values” to approximate joint localizations</b>	3
<b>III. Joint and sequential measurements</b>	5
A. Joint measurements	6
B. Sequential measurements	6
C. On measures of inaccuracy	7
1. Standard measures of error and disturbance	7
2. Geometric measure of approximation and disturbance	8
3. Error bars	8
<b>IV. From “no joint measurements” to approximate joint measurements</b>	9
A. Commuting functions of position and momentum	9
B. Uncertainty relations for covariant approximations of position and momentum	9
C. Uncertainty relations for general approximate joint measurements	11
<b>V. From “no measurement without disturbance” to sequential joint measurements</b>	12
<b>VI. Illustration: the Arthurs-Kelly model</b>	12
<b>VII. On experimental implementations and tests of the uncertainty principle</b>	14
A. Tests of preparation uncertainty relations	14
B. On implementations of joint and sequential measurements	15
C. On some alleged violations of the uncertainty principle	15
<b>VIII. Conclusion</b>	17
<b>References</b>	18

---

\*Electronic address: pb516@york.ac.uk

<sup>†</sup>Electronic address: teiko.heinonen@utu.fi

<sup>‡</sup>Electronic address: pekka.lahti@utu.fi

“It is the theory which decides what can be observed.”  
(Albert Einstein according to Werner Heisenberg (1977))

## I. INTRODUCTION

It seems to be no exaggeration to say that Heisenberg’s uncertainty principle, symbolized by the famous inequality for position and momentum,

$$\Delta q \cdot \Delta p \gtrsim \hbar, \quad (1)$$

epitomizes quantum physics, even in the eyes of the scientifically informed public. Nevertheless, still now, 80 years after its inception, there is no general consensus over the scope and validity of this principle. The aim of this article is to demonstrate that based on recent work, it has finally become possible to elucidate the full content of the uncertainty principle in precise terms. This will be done for the prime example of the position-momentum pair.

The uncertainty principle is usually described, rather vaguely, as comprising one or more of the following *no-go* statements, each of which will be made precise below:

- (A) *It is impossible to prepare states in which position and momentum are simultaneously arbitrarily well localized.*
- (B) *It is impossible to measure simultaneously position and momentum.*
- (C) *It is impossible to measure position without disturbing momentum, and vice versa.*

The negative characterization of the uncertainty principle as a limitations of quantum preparations and measurements has led to the widespread view that this principle is nothing but a formal expression of the principle of complementarity. This limited perspective has led some authors to question the fundamental status of the relation (1) (Scully *et al.*, 1991).

Here we will show that the uncertainty principle does have an independent content and role that in our view has not yet been duly recognized. In fact, instead of resigning to accepting the negative verdicts (A), (B), (C), it is possible to adopt a positive perspective on the underlying questions of joint preparation and measurement: according to the uncertainty principle, the qualitative relationship of a strict mutual exclusiveness of sharp preparations or measurements of position and momentum is complemented with a quantitative statement on a trade-off between competing degrees of the concentration of the distributions of these observables in state preparations or between the accuracies in joint measurements. Similarly, it turns out that the extent of the disturbance of (say) momentum in a measurement of position can be controlled if a limitation in the accuracy of that measurement is accepted.

Only if taken together, the statements (A), (B), (C) and their positive counterparts can be said to exhaust the content of the uncertainty principle for position and momentum. It also follows that the uncertainty principle comprises three conceptually distinct types of uncertainty relations.

We will give a systematic exposition of these three faces of the uncertainty principle, with an emphasis on elucidating its positive role. After a brief discussion of the well known uncertainty relation for preparations, we focus on the less well established *measurement* uncertainty relations, the formulation of which requires a careful discussion of joint measurements, measurement accuracy and disturbance.

We present a fundamental result, proved only very recently, which constitutes the first rigorous demonstration of the fact that the uncertainty relation for measurement inaccuracies is not only a sufficient but also a necessary condition for the possibility of approximate joint measurements of position and momentum.

Finally, we discuss some models and proposed realizations of joint measurements of position and momentum and address the question of possible experimental tests of the uncertainty principle.

The idea of the uncertainty principle ensuring the positive possibility of joint albeit imprecise measurements, which is rather latent in Heisenberg’s works<sup>1</sup> has been made fully explicit and brought to our attention by his former student Peter Mittelstaedt, our teacher and mentor, to whom we dedicate this treatise.

---

<sup>1</sup> A judicious reading of Heisenberg’s seminal paper of 1927 (Heisenberg, 1927) shows that both the double role and the three variants of the uncertainty principle discussed here are already manifest, if only expressed rather vaguely. In fact, in the abstract, Heisenberg immediately refers to limitations of joint measurements; later in the paper, he links this with a statement of the uncertainty relation for the widths of a Gaussian wave function and its Fourier transform; finally he gives illustrations by means of thought experiments in which the idea of mutual disturbance is prominent.

## II. FROM “NO JOINT SHARP VALUES” TO APPROXIMATE JOINT LOCALIZATIONS

Throughout the paper, we will only consider the case of a spin zero quantum system in one spatial dimension, represented by the Hilbert space  $\mathcal{H} = L^2(\mathbb{R})$ . The states of the system are described by positive trace one operators  $\rho$  on  $\mathcal{H}$ , the pure states being given as the one-dimensional projections. We occasionally write  $|\psi\rangle\langle\psi|$  for the pure state in question, and we call unit vectors  $\psi \in \mathcal{H}$  *vector states*. We denote the set of vector states by  $\mathcal{H}_1$ .

The position and momentum of the system are represented as the Schrödinger pair of operators  $\hat{Q}, \hat{P}$ , where  $\hat{Q}\psi(x) = x\psi(x)$ ,  $\hat{P}\psi(x) = -i\hbar\psi'(x)$ . We denote their spectral measures by the letters  $\mathbf{Q}$  and  $\mathbf{P}$ , respectively, and recall that they are Fourier-Plancherel connected. The probability of obtaining the value of position in a (Borel) subset  $X$  of  $\mathbb{R}$  on measurement in a vector state  $\psi$  is then given by the formula  $\mathbf{p}_\psi^{\mathbf{Q}}(X) = \langle\psi|\mathbf{Q}(X)\psi\rangle = \int_X |\psi(x)|^2 dx$ . Similarly, the probability of obtaining the value of momentum in a (Borel) set  $Y$  on measurement in a vector state  $\psi$  is given by  $\mathbf{p}_\psi^{\mathbf{P}}(Y) := \int_Y |\hat{\psi}(p)|^2 dp$ , where  $\hat{\psi}$  is the Fourier-Plancherel transform of  $\psi$ .

In the formalization of all three types of uncertainty relations, we will use two different measures of the width, or degree of concentration, of a probability distribution. These are the standard deviation and the overall width. The *standard deviations* of position and momentum in a state  $\psi$  are

$$\Delta(\mathbf{Q}, \psi) := \left( \langle\psi|\hat{Q}^2\psi\rangle - \langle\psi|\hat{Q}\psi\rangle^2 \right)^{1/2}, \quad \Delta(\mathbf{P}, \psi) := \left( \langle\psi|\hat{P}^2\psi\rangle - \langle\psi|\hat{P}\psi\rangle^2 \right)^{1/2}. \quad (2)$$

The *overall width* of a probability measure  $\mathbf{p}$  on  $\mathbb{R}$  is defined, for given  $\varepsilon \in (0, 1)$ , as the smallest interval length required to have probability greater or equal to  $1 - \varepsilon$ ; thus:

$$W_\varepsilon(\mathbf{p}) := \inf_X \{ |X| \mid \mathbf{p}(X) \geq 1 - \varepsilon \}, \quad (3)$$

where  $X$  run through all intervals in  $\mathbb{R}$ . For the overall widths of the position and momentum distribution in a vector state  $\psi$  we will use the notation

$$W_{\varepsilon_1}(\mathbf{Q}, \psi) := W_{\varepsilon_1}(\mathbf{p}_\psi^{\mathbf{Q}}), \quad W_{\varepsilon_2}(\mathbf{P}, \psi) := W_{\varepsilon_2}(\mathbf{p}_\psi^{\mathbf{P}}). \quad (4)$$

The no-go statement (A) says, broadly speaking, that the distributions  $\mathbf{p}_\psi^{\mathbf{Q}}$  and  $\mathbf{p}_\psi^{\mathbf{P}}$  of position and momentum cannot simultaneously (i.e., in the same state  $\psi$ ) be arbitrarily sharply concentrated. To appreciate this fact, we note first that position and momentum, being continuous quantities, cannot be assigned absolutely sharp values since they have no eigenvalues. But both quantities can separately have arbitrarily sharply concentrated distributions. We discuss two different ways of formalizing this idea, in terms of standard deviations and overall widths. Each of these formalizations gives rise to a precise form of (A). We then proceed to complement the no-go statement (A) with descriptions of the positive possibilities of simultaneous *approximate localizations* of position and momentum.

The first formalization of arbitrarily sharp localizations makes use of the standard deviation of a distribution (stated here for position):

$$\text{for any } q_0 \in \mathbb{R} \text{ and any } \varepsilon > 0, \text{ there is a vector state } \psi \text{ such that } \langle\psi|\hat{Q}\psi\rangle = q_0 \text{ and } \Delta(\mathbf{Q}, \psi) < \varepsilon. \quad (5)$$

Thus there is no obstacle to concentrating the distributions of position or momentum arbitrarily sharply at any points  $q_0, p_0 \in \mathbb{R}$ , if these observables are considered separately, on different sets of states.

But the corresponding state preparation procedures are mutually exclusive; this is the operational content of the negative statement (A). What *is* positively possible if one considers both observables together in the same state will be described by an appropriate uncertainty relation. Property (5) gives rise to the following formalization of (A):

**Theorem 1** *For all states  $\psi$  and for any  $\varepsilon > 0$ ,*

$$\text{if } \Delta(\mathbf{Q}, \psi) < \varepsilon, \quad \text{then } \Delta(\mathbf{P}, \psi) > \hbar/2\varepsilon; \quad \text{and vice versa.} \quad (6)$$

This is a statement about the spreads of the position and momentum probability distributions in a given state: the sharper one is peaked, the wider the other must be. This limitation follows directly from the uncertainty relation for standard deviations, valid for all vector states  $\psi$ :

$$\Delta(\mathbf{Q}, \psi) \cdot \Delta(\mathbf{P}, \psi) \geq \frac{\hbar}{2}. \quad (7)$$

The vector states  $\eta_{a,b}(x) = (2a/\pi)^{1/4} e^{-(a+ib)x^2}$ ,  $a, b \in \mathbb{R}$ ,  $a > 0$ , give  $\Delta(\mathbf{Q}, \eta_{a,b})^2 = 1/4a$  and  $\Delta(\mathbf{P}, \eta_{a,b})^2 = \hbar^2(a^2 + b^2)/a$ , so that the following positive statement complementing the no-go Theorem 1 is obtained:

**Theorem 2** For all positive numbers  $\delta q, \delta p$  for which  $\delta q \cdot \delta p \geq \hbar/2$ , there is a state  $\psi$  such that  $\Delta(Q, \psi) = \delta q$  and  $\Delta(P, \psi) = \delta p$ .

The vector state  $\eta_{a,0}$  is a *minimal uncertainty state* in the sense that it gives  $\Delta(Q, \eta_{a,0}) \cdot \Delta(P, \eta_{a,0}) = \hbar/2$ . Every minimum uncertainty state is of the form  $e^{icx} \eta_{a,0}(x-d)$  for some  $c, d \in \mathbb{R}$ . Minimal uncertainty states have a number of distinctive properties (see, e.g., (Dym and McKean, 1972)). For instance, if  $\psi$  is a vector state which satisfies  $|\psi|^2 \leq |\eta_{a,0}|^2$  and  $|\widehat{\psi}|^2 \leq |\widehat{\eta_{a,0}}|^2$ , for some  $a$ , then  $\psi$  is a minimal uncertainty state.

We now turn to the second way of saying that position and momentum can separately be localized arbitrarily well (expressed again only for position):

$$\text{for any bounded interval } X \text{ (however small), there exists a vector state } \psi \text{ such that } p_{\psi}^Q(X) = 1. \quad (8)$$

The corresponding formalization of (A) then is given by the following theorem.

**Theorem 3** For all vector states  $\psi$  and all bounded intervals  $X, Y$ ,  $p_{\psi}^Q(X) = 1$  implies  $0 \neq p_{\psi}^P(Y) \neq 1$ , and vice versa.

This means that whenever the position is localized in a bounded interval then the momentum cannot be confined to any bounded interval (nor to its complement), and vice versa.

For any two bounded intervals  $X$  and  $Y$  and for any vector state  $\psi$ , Theorem 3 implies that  $p_{\psi}^Q(X) + p_{\psi}^P(Y) < 2$ . However, for any such intervals, one can construct a vector state  $\psi_0$  for which the sum of the probabilities  $p_{\psi_0}^Q(X)$  and  $p_{\psi_0}^P(Y)$  attains its maximum value. The precise statement is given in the following theorem, which can be regarded as a positive complement to Theorem 3.

**Theorem 4** For any vector state  $\psi$  and for any bounded intervals  $X$  and  $Y$ ,

$$p_{\psi}^Q(X) + p_{\psi}^P(Y) \leq 1 + \sqrt{a_0} < 2, \quad (9)$$

where  $a_0$  is the largest eigenvalue of the positive trace class operator  $Q(X)P(Y)Q(X)$ . There exists an optimizing vector state  $\varphi_0$  such that

$$p_{\varphi_0}^Q(X) + p_{\varphi_0}^P(Y) = 1 + \sqrt{a_0}. \quad (10)$$

This result follows from the work of Landau and Pollak (1961) and Lenard (1972) (for details, see (Lahti, 1986)).

We will say that position  $Q$  is *approximately localized* in an interval  $X$  for a given state  $\psi$  whenever  $p_{\psi}^Q(X) \geq 1 - \varepsilon$  for some (preferably small)  $\varepsilon$ ,  $0 < \varepsilon < 1$ , and similarly for momentum. Then Eq. (10) describes the maximum degree of approximate localization that can be achieved in any phase space cell of given size  $|X| \cdot |Y|$ .

The largest eigenvalue  $a_0$  is invariant under a scale transformation applied simultaneously to  $\hat{Q}$  and  $\hat{P}$ ; it is therefore a function of the product  $|X| \cdot |Y|$  of the interval lengths  $|X|$  and  $|Y|$ . A simple calculation gives  $\text{tr}[Q(X)P(Y)Q(X)] = |X| \cdot |Y| / (2\pi\hbar)$ , so that we obtain

$$|X| \cdot |Y| \geq 2\pi\hbar \cdot a_0. \quad (11)$$

If both position and momentum are approximately localized within  $X$  and  $Y$ , respectively, so that  $p_{\psi}^Q(X) \geq 1 - \varepsilon_1$  and  $p_{\psi}^P(Y) \geq 1 - \varepsilon_2$ , then due to inequality (9), one must have  $1 - \varepsilon_1 - \varepsilon_2 \leq \sqrt{a_0}$ , and then (11) implies:

$$|X| \cdot |Y| \geq 2\pi\hbar \cdot (1 - \varepsilon_1 - \varepsilon_2)^2 \quad \text{if } p_{\psi}^Q(X) \geq 1 - \varepsilon_1, \quad p_{\psi}^P(Y) \geq 1 - \varepsilon_2, \quad \text{and } \varepsilon_1 + \varepsilon_2 < 1. \quad (12)$$

It is convenient to express this uncertainty relation for approximate localization widths in terms of the overall widths:

$$W_{\varepsilon_1}(Q, \psi) \cdot W_{\varepsilon_2}(P, \psi) \geq 2\pi\hbar \cdot (1 - \varepsilon_1 - \varepsilon_2)^2 \quad \text{if } \varepsilon_1 + \varepsilon_2 < 1. \quad (13)$$

If  $\varepsilon_1 + \varepsilon_2 \geq 1$ , then the product of widths has no positive lower bound (Landau and Pollak, 1961). In the case  $\varepsilon_1 + \varepsilon_2 < 1$ , the inequality is tight in the sense that even for fairly small values of  $\varepsilon_1, \varepsilon_2$ , the product of overall widths can be in the order of  $2\pi\hbar$ ; we quote a numerical example given in (Landau and Pollak, 1961): if  $\varepsilon_1 = \varepsilon_2 = .01$ , then  $|X| \cdot |Y|$  can still be as small as  $6.25 \times (2\pi\hbar)$ .

We will make repeated use of this uncertainty relation; since we are only interested in ‘‘small’’ values of  $\varepsilon_1, \varepsilon_2$ , we will assume these to be less than  $1/2$ ; then the condition  $\varepsilon_1 + \varepsilon_2 < 1$  is fulfilled and need not be stated explicitly.

Several other measures of uncertainty have been introduced to analyze the degree of (approximate) localizability of position and momentum distributions  $\mathbf{p}_\psi^Q$  and  $\mathbf{p}_\psi^P$ , ranging from extensive studies on the support properties of  $|\psi|^2$  and  $|\widehat{\psi}|^2$  to various information theoretic (“entropic”) uncertainty relations. It is beyond the scope of this paper to review the vast body of literature on this topic. The interested reader may consult e.g. (Cowling and Price, 1984), (Folland and Sitaram, 1997) or (Busch *et al.*, 1997, Sect. V.4) for reviews and references.

To summarize: instead of leaving it at the negative statement that position and momentum *cannot be arbitrarily sharply localized in the same state*, the uncertainty relation for state preparations offers precise specifications of the extent to which these two observables *can simultaneously be approximately localized*.

### III. JOINT AND SEQUENTIAL MEASUREMENTS

In order to go beyond the no-go theorems of (B) and (C) and establish their positive complements, one needs to use the full-fledged apparatus of quantum mechanics. The general representation of observables as positive operator measures will be required to introduce viable notions of joint and sequential measurements and an appropriate quantification of measurement inaccuracy. Furthermore, some tools of measurement theory will be needed to describe and quantify the disturbance of one observable due to the measurement of another.

In the present context of the discussion of position and momentum and their joint measurements, observables will be described as positive operator measures  $X \mapsto E(X)$  on the (Borel) subsets of  $\mathbb{R}$  or  $\mathbb{R}^2$ . This means that the map  $X \mapsto \langle \psi | E(X) \psi \rangle =: \mathbf{p}_\psi^E(X)$  is a normalized probability measure for every vector state  $\psi$ . An observable  $E$  will be called *sharp* if it is a spectral measure, that is, if all of the effects  $E(X)$  are projections.

For an observable  $E$  on  $\mathbb{R}$ , we will make use of the notation  $E[1]$ ,  $E[2]$  for the first and the second moment operators, defined (weakly) as  $E[k] := \int x^k E(dx)$  ( $k = 1, 2$ ).<sup>2</sup> We let  $\Delta(E, \psi)$  denote the standard deviation of  $\mathbf{p}_\psi^E$ ,

$$\Delta(E, \psi)^2 := \int_{-\infty}^{\infty} \left( x - \int_{-\infty}^{\infty} x' \mathbf{p}_\psi^E(dx') \right)^2 \mathbf{p}_\psi^E(dx) = \langle \psi | E[2] \psi \rangle - \langle \psi | E[1] \psi \rangle^2. \quad (14)$$

It is a remarkable feature of an observable  $E$ , defined as a positive operator measure, that it need not be commutative; that is, it is not always the case that  $E(X_1)E(X_2) = E(X_2)E(X_1)$  for all sets  $X_1, X_2$ . This opens up the possibility of defining a notion of joint measurability for not necessarily commuting families of observables.

Indeed, it will become evident below that in the set of noncommuting pairs of observables, the jointly measurable ones are necessarily unsharp. It is to be expected intuitively that the degree of mutual noncommutativity determines the necessary degree of unsharpness required to allow a joint measurement. Here we present two ways of indicating the inherent unsharpness of an observable  $E$  on  $\mathbb{R}$ .

We define the *intrinsic noise operator* of  $E$  as

$$N_i(E) := E[2] - E[1]^2. \quad (15)$$

This is a positive operator. If  $E[1]$  is selfadjoint, then the intrinsic noise  $N_i(E)$  is zero exactly when  $E$  is a sharp observable (Kiukas *et al.*, 2006b, Theorem 5). A measure  $\nu_i(E; \psi)$  of *intrinsic noise* is then given by the expectation value of the intrinsic noise operator (for all vector states  $\psi$  for which this expression is well defined):

$$\nu_i(E; \psi) := \langle \psi | N_i(E) \psi \rangle. \quad (16)$$

The *overall intrinsic noise* is defined as

$$\nu_i(E) := \sup_{\psi \in \mathcal{H}_1} \nu_i(E; \psi). \quad (17)$$

The next definition applies to observables  $E$  on  $\mathbb{R}$  whose support is  $\mathbb{R}$ .<sup>3</sup> The *resolution width* of  $E$  (at confidence level  $1 - \varepsilon$ ) is (Carmeli *et al.*, 2006)

$$\gamma_\varepsilon(E) := \inf\{d > 0 \mid \text{for all } x \in \mathbb{R} \text{ there exists } \psi \in \mathcal{H}_1 \text{ with } \mathbf{p}_\psi^E(J_{x,d}) \geq 1 - \varepsilon\}. \quad (18)$$

<sup>2</sup> It is important to bear in mind that the domain of the symmetric operator  $E[k]$  is not necessarily dense; it consists of all vectors  $\psi \in \mathcal{H}$  for which the function  $x \mapsto x^k$  is integrable w.r.t. the complex measure  $X \mapsto \langle \psi | E(X) \psi \rangle$  for all  $\psi \in \mathcal{H}$ . Here and in subsequent formulas it is understood that the expectations of unbounded operators are only well defined for appropriate subsets of states.

<sup>3</sup> This means that for every interval  $J$  there is a vector state  $\psi$  such that  $\mathbf{p}_\psi^E(J) \neq 0$ .

Here  $J_{x;d}$  denotes the interval  $[x - \frac{d}{2}, x + \frac{d}{2}]$ .

We note that positive resolution width is a certain indicator that the observable  $E$  is unsharp. This measure describes the possibilities of concentrating the probability distributions to a fixed confidence level across all intervals. However, the requirement of vanishing resolution width does not single out sharp observables (Carmeli *et al.*, 2006).

### A. Joint measurements

Two observables  $E_1$  and  $E_2$  on  $\mathbb{R}$  are called *jointly measurable* if there is an observable  $M$  on  $\mathbb{R}^2$  such that

$$E_1(X) = M(X \times \mathbb{R}), \quad E_2(Y) = M(\mathbb{R} \times Y) \quad (19)$$

for all (Borel) sets  $X, Y$ . Then  $E_1$  and  $E_2$  are the marginal observables  $M_1$  and  $M_2$  of the joint observable  $M$ . If either  $E_1$  or  $E_2$  is a sharp observable, then they are jointly measurable exactly when they commute mutually. In that case, the unique joint observable  $M$  is determined by  $M(X \times Y) = E_1(X)E_2(Y)$ . In general, the mutual commutativity of  $E_1$  and  $E_2$  is not a necessary (although still a sufficient) condition for their joint measurability.

The above notion of joint measurability is fully supported by the quantum theory of measurement, which ensures that  $E_1$  and  $E_2$  are jointly measurable exactly when there is a measurement scheme which measures both  $E_1$  and  $E_2$  (Busch *et al.*, 1996).

Considering that the (sharp) position and momentum observables  $Q$  and  $P$  do not commute with each other, we recover immediately the well-known fact that these observables have no joint observable, that is, they are not jointly measurable. This is a precise formulation of the no-go statement (B).

In preparation of developing a positive complementation to (B), we give an outline of the notion that sharp position and momentum may be jointly measurable in an approximate sense. While there is no observable on phase space whose marginals coincide with  $Q$  and  $P$ , one can explore the idea that there may be observables  $M$  on  $\mathbb{R}^2$  whose marginals  $M_1$  and  $M_2$  are *approximations* (in some suitably defined sense) of  $Q$  and  $P$ , respectively. Such an  $M$  will be called an *approximate joint observable* for  $Q$  and  $P$ . An appropriate quantification of the differences between  $M_1$  and  $Q$  and between  $M_2$  and  $P$  may serve as a measure of the (*in*)accuracy of the joint approximate measurement represented by  $M$ .

### B. Sequential measurements

In order to analyze measurements of two observables  $E_1$  and  $E_2$  performed in immediate succession, it is necessary to take into account the influence of the first measurement on the object system. The tool to describe the state changes due to a measurement is provided by the notion of an *instrument*; see, for instance, (Davies, 1976, Chapter 4).

Let  $\mathcal{I}_1$  be the instrument associated with a measurement of  $E_1$ , that is,  $\mathcal{I}_1$  determines the probability  $\text{tr}[\rho E_1(X)] = \text{tr}[\mathcal{I}_1(X)(\rho)]$  for every state  $\rho$  and set  $X$ . The number  $\text{tr}[\mathcal{I}_1(X)(\rho)E_2(Y)]$  is the sequential joint probability that the measurement of  $E_1$ , performed on the system in state  $\rho$ , gives a result in  $X$  and a subsequent measurement of  $E_2$  leads to a result in  $Y$ . Using the dual instrument  $\mathcal{I}_1^*$  (i.e., the Heisenberg picture), this probability can be written as  $\text{tr}[\rho \mathcal{I}_1^*(X)(E_2(Y))]$ . The map  $(X, Y) \mapsto \text{tr}[\rho \mathcal{I}_1^*(X)(E_2(Y))]$  is a probability bimeasure and therefore extends uniquely to a joint probability for each  $\rho$ , defining thus a unique joint observable  $M$  on  $\mathbb{R}^2$  via

$$M(X \times Y) = \mathcal{I}_1^*(X)(E_2(Y)). \quad (20)$$

Its marginal observables are

$$M_1(X) = \mathcal{I}_1^*(X)(E_2(\mathbb{R})) = E_1(X), \quad (21)$$

$$M_2(Y) = \mathcal{I}_1^*(\mathbb{R})(E_2(Y)) =: E_2'(Y). \quad (22)$$

Thus the first marginal is the first-measured observable  $E_1$  and the second marginal is a distorted version  $E_2'$  of  $E_2$ .

This general consideration shows that one must expect that a measurement of an observable  $E_1$  will *disturb* (the distribution of) another observable  $E_2$ . In fact, it is a fundamental theorem of the quantum theory of measurement that there is no nontrivial measurement without *some* state changes. In other words, if a measurement leaves *all* states unchanged, then its statistics will be the same for all states; in this sense there is no information gain without some disturbance.

If the first-measured observable  $E_1$  is sharp, then the distorted effects  $\mathcal{I}_1^*(\mathbb{R})(E_2(Y))$  must commute with  $E_1(X)$  for all  $X, Y$ , whatever the second observable  $E_2$  is. Thus, if we consider a sequential measurement of the (sharp) position

and momentum observables  $Q$  and  $P$  as an attempted joint measurement, we see that such an attempt is bound to fail. If (say) one first measures position  $Q$ , with an instrument  $\mathcal{I}_Q$ , then all distorted momentum effects  $P'(Y) := \mathcal{I}_Q^*(Y)$  are functions of the position operator  $\hat{Q}$ . In this sense, a measurement of (sharp) position completely destroys any information about the momentum distribution in the input state. This result formalizes the no-go statement (C).

The formulation of a positive complement to (C) is based on the idea that one may be able to control and limit the disturbance due to a measurement of  $Q$ , by measuring an observable  $Q'$  which is an approximation (in some sense) to  $Q$ . One can then hope to achieve that the distorted momentum  $P'$  is an approximation (in some sense) to  $P$ . We note that this amounts to defining a sequential joint observable  $M$  with marginals  $M_1 = Q'$  and  $M_2 = P'$ . Any appropriate quantification of the difference between  $M_1$  and  $Q$  is a measure of the inaccuracy of the first (approximate) position measurement; similarly any appropriate quantification of the difference between  $M_2$  and  $P$  is a measure of the disturbance of the momentum due to the position measurement.

In this way the problem of defining measures of the disturbance of (say) momentum due to a measurement of position has been reduced to defining the inaccuracy of the second marginal of a sequential joint measurement of first position and then momentum.

### C. On measures of inaccuracy

The above discussion shows that it is the noncommutativity of observables such as position and momentum which forces one to allow inaccuracies if one attempts to make an approximate joint measurement of these observables. This shows clearly that the required inaccuracies are of quantum-mechanical origin, which will also become manifest in the models of approximate position measurements and phase space measurements presented below. With this observation as proviso, we believe that it is acceptable to use the classical terms of measurement inaccuracy and error, particularly because their operational definitions are essentially the same as in a classical measurement context.

In fact, every measurement, whether classical or quantum, is subject to *noise*, which results in a deviation of the actually measured observable  $E_1$  from that intended to be measured,  $E$ . We will refer to this deviation and any measure of it as *error* or *inaccuracy*. In general there can be systematic errors, or *bias*, leading to a shift of the mean values, and random errors, resulting in a broadening of the distributions. Any measure of measurement noise should be *operationally significant* in the sense that it should be determined by the probability distributions  $\mathfrak{p}_\psi^E$  and  $\mathfrak{p}_\psi^{E_1}$ .

In the following we will discuss three different approaches to quantifying measurement inaccuracy.

#### 1. Standard measures of error and disturbance

Classical statistical analysis suggests the use of moments of probability distributions for the quantification of error and disturbance in measurements. Thus, the standard approach found in the literature of defining a measure of error is in terms of the average deviation of the value of a readout observable of the measuring apparatus from the value of the observable to be measured approximately. If these observables are represented as selfadjoint operators  $Z$  and  $A$  (acting on the apparatus and the object Hilbert spaces), respectively, this *standard error* measure is given as the root mean square

$$\epsilon(Z, A, \psi) := \langle U(\psi \otimes \Psi) | (Z - A)^2 U(\psi \otimes \Psi) \rangle^{1/2}, \quad (23)$$

where  $U$  is the unitary map modelling the measuring interaction and  $\Psi$  is the initial state of the apparatus. This measure of error has been studied in recent years in the foundational context, for example, by Appleby (1998a,b), Hall (2004) and Ozawa (2004).

If we denote by  $E$  the observable actually measured by the given scheme, we define the *relative noise* operator,

$$N_r(E, A) := E[1] - A; \quad (24)$$

the standard error can then be rewritten as (Ozawa, 2004)

$$\epsilon(E, A; \psi)^2 = \langle \psi | (E[1] - A)^2 \psi \rangle + \langle \psi | (E[2] - E[1]^2) \psi \rangle = \langle \psi | N_r(E, A)^2 \psi \rangle + \langle \psi | N_i(E) \psi \rangle \quad (25)$$

(for any vector state  $\psi$  for which the expressions are well-defined). We note that  $\epsilon(E, A; \psi) = 0$  for all  $\psi$  exactly when  $E$  is sharp and  $E[1] = A$ . The relative noise term cannot, in general, be determined from the statistics of measurements of  $E$  and  $A$  alone, so that the standard error measure  $\epsilon(E, A; \psi)$  does not always satisfy the requirement of operational significance (Busch *et al.*, 2004). However, we will encounter important cases where this quantity does turn out to be operationally well defined.

The standard error is a state-dependent quantity. This stands in contrast to the fact that estimates of errors obtained in a calibration process are meant to be applicable to a range of states since in a typical measurement the state is unknown to begin with. In order to obtain state independent measures, we define the *global standard error* of an observable  $E$  relative to  $A$  as

$$\epsilon(E, A) := \sup_{\psi \in \mathcal{H}_1} \epsilon(E, A; \psi). \quad (26)$$

We will say that  $E$  is a *standard approximation* to  $A$  if  $E$  has finite global standard error relative to  $A$ . This definition provides a possible criterion for selecting joint or sequential measurements schemes as approximate joint measurements of  $Q$  and  $P$ ; but it is not always possible to verify this criterion if the standard error fails to be operationally significant.

## 2. Geometric measure of approximation and disturbance

Following the work of Werner (2004), we define a distance  $d(E_1, E_2)$  on the set of observables on  $\mathbb{R}$ .

We first recall that for any bounded measurable function  $h : \mathbb{R} \rightarrow \mathbb{R}$ , the integral  $\int_{\mathbb{R}} h dE$  defines (in the weak sense) a bounded selfadjoint operator, which we denote by  $E[h]$ . Thus, for any vector state  $\psi$  the number  $\langle \psi | E[h] \psi \rangle = \int_{\mathbb{R}} h d\mathbf{p}_{\psi}^E$  is well-defined.

Denoting by  $\Lambda$  the set of bounded measurable functions  $h : \mathbb{R} \rightarrow \mathbb{R}$  for which  $|h(x) - h(y)| \leq |x - y|$ , the distance between the observables  $E_1$  and  $E_2$  is defined as

$$d(E_1, E_2) := \sup_{\psi \in \mathcal{H}_1} \sup_{h \in \Lambda} |\langle \psi | (E_1[h] - E_2[h]) \psi \rangle|. \quad (27)$$

This measure is operationally significant, using only properties of the distributions to be compared. Furthermore, it is a global measure in that it takes into account the largest possible deviations of the expectations  $\langle \psi | E_1[h] \psi \rangle$  and  $\langle \psi | E_2[h] \psi \rangle$ . It gives a geometrically appealing quantification of how well a given observable can be approximated by other observables.

We will say that an observable  $E_1$  is a *geometric approximation* to  $E_2$  if  $d(E_1, E_2) < \infty$ . We shall apply this condition of finite distance as a criterion for a joint or sequential measurement scheme to define an approximate joint measurement of  $Q$  and  $P$ . It is not clear whether this criterion is practical since the distance is not related in any obvious way to concepts of measurement inaccuracy commonly applied in an experimental context.

## 3. Error bars

We now present a definition of measurement inaccuracy in terms of likely error intervals that follows most closely the usual practice of calibrating measuring instruments. In the process of calibration of a measurement scheme, one seeks to obtain estimates of the likely error and perhaps also the degree of disturbance that the scheme contains. To estimate the error, one tests the device by applying it to a sufficiently large family of input states in which the observable one wishes to measure with this setup has fairly sharp values. The error is then characterized as an overall measure of the bias and the width of the output distribution across a range of input values. Error bars give the minimal average interval lengths that one has to allow to contain all output values with a given confidence level.

For each  $\varepsilon \in (0, 1)$ , we say that an observable  $E_1$  is an  $\varepsilon$ -*approximation* to a sharp observable  $E$  if for all  $\delta > 0$  there is a positive number  $w < \infty$  such that for all  $x \in \mathbb{R}$ ,  $\psi \in \mathcal{H}_1$ , the condition  $\mathbf{p}_{\psi}^E(J_{x;\delta}) = 1$  implies that  $\mathbf{p}_{\psi}^{E_1}(J_{x,w}) \geq 1 - \varepsilon$ . The infimum of all such  $w$  will be called the *inaccuracy* of  $E_1$  with respect to  $E$  and will be denoted  $\mathcal{W}_{\varepsilon,\delta}(E_1, E)$ . Thus,

$$\mathcal{W}_{\varepsilon,\delta}(E_1, E) := \inf\{w \mid \text{for all } x \in \mathbb{R} : \mathbf{p}_{\psi}^{E_1}(J_{x,w}) \geq 1 - \varepsilon \text{ for all } \psi \in \mathcal{H}_1 \text{ with } \mathbf{p}_{\psi}^E(J_{x;\delta}) = 1\}. \quad (28)$$

The inaccuracy describes the range within which the input values can be inferred from the output distributions, with confidence level  $1 - \varepsilon$ , given initial localizations within  $\delta$ . We note that the inaccuracy is an increasing function of  $\delta$ , so that we can define the *error bar width*<sup>4</sup> of  $E_1$  relative to  $E$ :

$$\mathcal{W}_{\varepsilon}(E_1, E) := \inf_{\delta} \mathcal{W}_{\varepsilon,\delta}(E_1, E) = \lim_{\delta \rightarrow 0} \mathcal{W}_{\varepsilon,\delta}(E_1, E). \quad (29)$$

---

<sup>4</sup> This definition and all subsequent results based on it can be found in (Busch and Pearson, 2006).

If  $\mathcal{W}_\varepsilon(E_1, E)$  is finite for all  $\varepsilon \in (0, \frac{1}{2})$ , we will say that  $E_1$  approximates  $E$  in the sense of *finite error bars*. We note that the finiteness of either  $\epsilon(E_1, E)$  or  $d(E_1, E)$  implies the finiteness of  $\mathcal{W}_\varepsilon(E_1, E)$ . Therefore, among the three measures of inaccuracy, the condition of finite error bars gives the most general criterion for selecting approximations of  $Q$  and  $P$ .

#### IV. FROM “NO JOINT MEASUREMENTS” TO APPROXIMATE JOINT MEASUREMENTS

Position  $Q$  and momentum  $P$  have no joint observable, they cannot be measured together. However, one may ask for an approximate joint measurement, that is, for an observable  $M$  on  $\mathbb{R}^2$  such that the marginals  $M_1$  and  $M_2$  are appropriate approximations of  $Q$  and  $P$ . In this section we study two important cases and then consider the general situation.

##### A. Commuting functions of position and momentum

The first approach is related with the fact that although  $Q$  and  $P$  are noncommutative, they do have commuting spectral projections. Indeed, let  $Q^g$  be a function of  $Q$ , that is,  $Q^g(X) = Q(g^{-1}(X))$  for all  $X \in \mathcal{B}(\mathbb{R})$ , with  $g : \mathbb{R} \rightarrow \mathbb{R}$  being a (Borel) function. Similarly, let  $P^h$  be a function of momentum. The associated operators are  $g(\hat{Q})$  and  $h(\hat{P})$ . The following result, proved by Busch *et al.* (1987, Theorem 1) and in a more general setting by Ylinen (1989), characterizes the functions  $g$  and  $h$  for which  $Q^g(X)P^h(Y) = P^h(Y)Q^g(X)$  for all  $X$  and  $Y$ .

**Theorem 5** *Let  $g$  and  $h$  be essentially bounded Borel functions such that neither  $g(\hat{Q})$  nor  $h(\hat{P})$  is a constant operator. The functions  $Q^g$  of position and  $P^h$  of momentum commute if and only if  $g$  and  $h$  are both periodic with minimal positive periods  $a, b$  satisfying  $\frac{2\pi}{ab} \in \mathbb{N}$ .*

If  $Q^g$  and  $P^h$  are commuting observables, then they have the joint observable  $M$ , with  $M(X \times Y) = Q^g(X)P^h(Y)$ , meaning that  $Q^g$  and  $P^h$  can be measured jointly. The price for this restricted form of joint measurability of position and momentum as given by Theorem 5 is that they are to be coarse-grained by periodic functions  $g$  and  $h$  with appropriately related minimal periods  $a, b$ .

The functions  $g$  and  $h$  can be chosen as characteristic functions of appropriate periodic sets. This allows one to model a situation known in solid state physics, where an electron in a crystal can be confined arbitrarily closely to the atoms while at the same time its momentum is localized arbitrarily closely to the reciprocal lattice points.

Simultaneous localization of position and momentum in periodic sets thus constitutes a sharp joint measurement of functions of these observables. However, bounded functions  $Q^f$  of  $Q$  provide only very bad approximations to  $Q$  since  $\epsilon(Q^f, Q)$ ,  $d(Q^f, Q)$  and  $\mathcal{W}_{\varepsilon, \delta}(Q^f, Q)$  are all infinite. One also loses the characteristic covariance properties of position and momentum.

##### B. Uncertainty relations for covariant approximations of position and momentum

Next we will discuss approximate joint measurements of position and momentum based on smearings of these observables by means of convolutions.

Let  $\mu, \nu$  be probability measures on  $\mathbb{R}$ . We define observables  $Q_\mu, P_\nu$  via

$$Q_\mu(X) = \int_{\mathbb{R}} \mu(X - q) Q(dq), \quad P_\nu(Y) = \int_{\mathbb{R}} \nu(Y - p) P(dp). \quad (30)$$

These observables have the same characteristic covariance properties as  $Q, P$  and they are approximations in the sense that they have finite error bar widths relative to  $Q, P$ . Hence we call them approximate position and momentum.

For given  $Q_\mu, P_\nu$  we ask under what conditions they are jointly measurable, that is, there is an observable  $M$  on  $\mathbb{R}^2$  such that  $M_1 = Q_\mu$  and  $M_2 = P_\nu$ . In order to answer this question, we need to introduce the notion of covariant phase space observables.

Covariance is defined with respect to a unitary (projective) representation of phase space translations in terms of the Weyl operators, defined for any phase space point  $(q, p) \in \mathbb{R}^2$  via  $W(q, p) = e^{\frac{i}{\hbar} qP} e^{-\frac{i}{\hbar} pQ}$ . An observable  $M$  on  $\mathbb{R}^2$  is a *covariant phase space observable* if

$$W(q, p)M(Z)W(q, p)^* = M(Z + (q, p)) \quad (31)$$

for all  $Z$ . It is known<sup>5</sup> that each such observable is of the form  $G^T$ , where

$$G^T(Z) = \frac{1}{2\pi\hbar} \int_Z W(q,p)TW(q,p)^* dqdp, \quad (32)$$

and  $T$  is a unique positive operator of trace one. The marginal observables  $G_1^T$  and  $G_2^T$  are of the form (30) where  $\mu = \mu_T = \mathfrak{p}_{\Pi T \Pi^*}^{\mathbf{Q}}$  and  $\nu = \nu_T = \mathfrak{p}_{\Pi T \Pi^*}^{\mathbf{P}}$  and  $\Pi$  is the parity operator,  $\Pi\psi(x) = \psi(-x)$ .

Our question is answered by the following fundamental theorem which was proven in the present form by Carmeli *et al.* (2004, Proposition 7) as a direct development of the work of Werner (2004).

**Theorem 6** *An approximate position  $\mathbf{Q}_\mu$  and an approximate momentum  $\mathbf{P}_\nu$  are jointly measurable if and only if they have a covariant joint observable  $G^T$ . This is the case exactly when there is a positive operator  $T$  of trace equal to 1 such that  $\mu = \mu_T$ ,  $\nu = \nu_T$ .*

We next quantify the necessary trade-off in the quality of the approximations of  $\mathbf{Q}, \mathbf{P}$  by  $G_1^T, G_2^T$ , using the three measures of inaccuracy introduced above. First we state uncertainty relations for the measures of intrinsic unsharpness. The probability distributions associated with these marginals of  $G^T$  are the convolutions  $\mathfrak{p}_\psi^{\mathbf{Q}} * \mu_T$  and  $\mathfrak{p}_\psi^{\mathbf{P}} * \nu_T$ ; this indicates that statistically independent noise is added to the distributions of sharp position and momentum. In accordance with this fact, the standard deviations are obtained via the sums of variances,

$$\Delta(G_1^T, \psi)^2 = \Delta(\mathbf{Q}, \psi)^2 + \Delta(\mu_T)^2, \quad \Delta(G_2^T, \psi)^2 = \Delta(\mathbf{P}, \psi)^2 + \Delta(\nu_T)^2. \quad (33)$$

The noise interpretation is confirmed by a determination of the intrinsic noise operators of  $G_1^T, G_2^T$ , which are well-defined whenever the operator  $T$  is such that  $\hat{Q}^2\sqrt{T}$  and  $\hat{P}^2\sqrt{T}$  are Hilbert-Schmidt operators (Kiukas *et al.*, 2006b, Theorem 4); in that case one obtains

$$N_i(G_1^T) = \Delta(\mu_T)^2 I, \quad N_i(G_2^T) = \Delta(\nu_T)^2 I. \quad (34)$$

Since  $T$ , and with it  $\Pi T \Pi^*$ , has the properties of a state, the uncertainty relation (7) applies to the probability measures  $\mu_T$  and  $\nu_T$ , giving the following uncertainty relation for intrinsic noise, valid for any  $G^T$ :

$$\nu_i(G_1^T) \cdot \nu_i(G_2^T) = \Delta(\mu_T) \cdot \Delta(\nu_T) \geq \frac{\hbar}{2}. \quad (35)$$

Equations (33) and (35) yield the following version of state-preparation uncertainty relation with respect to  $G^T$ , which also reflects the presence of the intrinsic noise:

$$\Delta(G_1^T, \psi) \cdot \Delta(G_2^T, \psi) \geq \hbar. \quad (36)$$

The resolution widths of  $G_1^T, G_2^T$  are given by

$$\gamma_{\varepsilon_1}(G_1^T) = W_{\varepsilon_1}(\mu_T), \quad \gamma_{\varepsilon_2}(G_2^T) = W_{\varepsilon_2}(\nu_T), \quad (37)$$

so that the uncertainty relation for overall widths then entails:

$$\gamma_{\varepsilon_1}(G_1^T) \cdot \gamma_{\varepsilon_2}(G_2^T) = W_{\varepsilon_1}(\mu_T) \cdot W_{\varepsilon_2}(\nu_T) \geq 2\pi\hbar \cdot (1 - \varepsilon_1 - \varepsilon_2)^2. \quad (38)$$

The standard errors are

$$\epsilon(G_1^T, \mathbf{Q}; \psi)^2 = (\mu_T[1])^2 + \Delta(\mu_T)^2, \quad \epsilon(G_2^T, \mathbf{P}; \psi)^2 = (\nu_T[1])^2 + \Delta(\nu_T)^2, \quad (39)$$

and the inequality

$$\epsilon(G_1^T, \mathbf{Q}) \cdot \epsilon(G_2^T, \mathbf{P}) \geq \frac{\hbar}{2} \quad (40)$$

holds as an immediate consequence of the noise uncertainty relation (35).

---

<sup>5</sup> This result is due to Holevo (1979) and Werner (1984). Alternative proofs with different techniques were recently given by Cassinelli *et al.* (2003) and Kiukas *et al.* (2006a).

The distances of  $G_1^T, G_2^T$  from  $\mathbf{Q}, \mathbf{P}$  are

$$d(G_1^T, \mathbf{Q}) = \int |q| \mu_T(dq), \quad d(G_2^T, \mathbf{P}) = \int |p| \nu_T(dp), \quad (41)$$

and they satisfy the trade-off inequality

$$d(G_1^T, \mathbf{Q}) \cdot d(G_2^T, \mathbf{P}) \geq C\hbar, \quad (42)$$

where the value of the constant  $C$  can be numerically determined as  $C \approx 0.3047$  (Werner, 2004). There is a unique covariant joint observable  $G^T$  attaining the lower bound in (42), but the optimizing operator  $T = |\eta\rangle\langle\eta|$  is not given by the oscillator ground state (Werner, 2004, Section 3.2).

Finally considering the error bar widths of  $G_1^T, G_2^T$  relative to  $\mathbf{Q}, \mathbf{P}$ , one finds:

$$\mathcal{W}_{\varepsilon_1}(G_1^T, \mathbf{Q}) \geq W_{\varepsilon_1}(\mathbf{Q}, T), \quad \mathcal{W}_{\varepsilon_2}(G_2^T, \mathbf{P}) \geq W_{\varepsilon_2}(\mathbf{P}, T). \quad (43)$$

Therefore, (13) implies that

$$\mathcal{W}_{\varepsilon_1}(G_1^T, \mathbf{Q}) \cdot \mathcal{W}_{\varepsilon_2}(G_2^T, \mathbf{P}) \geq 2\pi\hbar \cdot (1 - \varepsilon_1 - \varepsilon_2)^2. \quad (44)$$

We note that error bar widths in this inequality are always finite, in contrast to the standard errors or distances, which are infinite for some  $G^T$ .

The existence of covariant phase space observables  $G^T$  establishes the positive complement to the no-go statement (B). We have given Heisenberg uncertainty relations for the *necessary* inaccuracies in the approximations of  $\mathbf{Q}, \mathbf{P}$  by means of the marginal observables  $G_1^T, G_2^T$ . For each pair of values of the inaccuracies allowed by these uncertainty relations there exists a  $G^T$  which realizes these values. This confirms the *sufficiency* of the inaccuracy relations for the existence of an approximate joint measurement of position and momentum, in the form of a covariant joint observable.

There is a (perhaps unexpected) reward for the positive attitude that led to the search for approximate joint measurements of position and momentum: the family of covariant phase space observables  $G^T$  contains the important class of *informationally complete* phase space observables. An example is given by the choice  $T = |\eta_{a,0}\rangle\langle\eta_{a,0}|$ .

### C. Uncertainty relations for general approximate joint measurements

While the uncertainty relations are necessary for the inaccuracies inherent in jointly measurable covariant approximations  $\mathbf{Q}_\mu$  and  $\mathbf{P}_\nu$ , there remains possibility that one can overcome the Heisenberg limit by some clever choice of non-covariant approximations of  $\mathbf{Q}$  and  $\mathbf{P}$ . Here we show that this possibility is ruled out. It follows that covariant phase space observables constitute the optimal class of approximate joint observables for position and momentum.

Let  $M$  be an observable on  $\mathbb{R}^2$ . It was shown by Werner (2004) that there is a covariant phase space observable  $G^T$  associated with  $M$  with the following property:  $d(M_1, \mathbf{Q}) \geq d(G_1^T, \mathbf{Q})$  and  $d(M_2, \mathbf{P}) \geq d(G_2^T, \mathbf{P})$ . The same kind of argument can be carried out in the case of the global standard error and the error bar width, so that the inequalities (40), (44) and (42) entail the universally valid Heisenberg uncertainty relations

$$\epsilon(M_1, \mathbf{Q}) \cdot \epsilon(M_2, \mathbf{P}) \geq \frac{\hbar}{2}, \quad (45)$$

$$d(M_1, \mathbf{Q}) \cdot d(M_2, \mathbf{P}) \geq C\hbar, \quad (46)$$

$$\mathcal{W}_{\varepsilon_1}(M_1, \mathbf{Q}) \cdot \mathcal{W}_{\varepsilon_2}(M_2, \mathbf{P}) \geq 2\pi\hbar \cdot (1 - \varepsilon_1 - \varepsilon_2)^2. \quad (47)$$

The same considerations leading to these inaccuracy relations can also be applied to obtain general trade-off relations for the intrinsic noise and resolution width of the marginals of an approximate joint observable of  $\mathbf{Q}, \mathbf{P}$ :

$$\nu_i(M_1) \cdot \nu_i(M_2) \geq \frac{\hbar}{2}, \quad (48)$$

$$\gamma_{\varepsilon_1}(M_1) \cdot \gamma_{\varepsilon_2}(M_2) \geq 2\pi\hbar \cdot (1 - \varepsilon_1 - \varepsilon_2)^2. \quad (49)$$

Considering now examples of noncovariant observables on phase space, we recall first that the commutative observable  $M$  on  $\mathbb{R}^2$  of subsection IV.A has marginals with infinite error bars. Here we give an example of an observable

$M$  on phase space which is not covariant but is still an approximate joint observable for  $\mathbf{Q}, \mathbf{P}$ . Let  $G^T$  be a covariant phase space observable and define  $M := G^T \circ \gamma^{-1}$ , where  $\gamma(q, p) := (\gamma_1(q), \gamma_2(p))$ . We assume that  $\gamma_1, \gamma_2$  are strictly increasing continuous functions such that  $\gamma_1(q) - q$  and  $\gamma_2(p) - p$  are bounded functions. Then it follows that the marginals  $M_1^\gamma = G_1^T \circ \gamma_1^{-1}$  and  $M_2^\gamma = G_2^T \circ \gamma_2^{-1}$  have finite error bars with respect to  $\mathbf{Q}, \mathbf{P}$ . If  $\gamma$  is a nonlinear function then  $M$  will not be covariant.

## V. FROM “NO MEASUREMENT WITHOUT DISTURBANCE” TO SEQUENTIAL JOINT MEASUREMENTS

As concluded in Subsection III.B there is no way to determine the (sharp) position and momentum observables in a sequential measurement. We show now that there are sequential measurements which are approximate simultaneous determinations of position and momentum. As discussed above, the inaccuracy of the second measurement defines an operational measure of the disturbance of momentum due to the first, approximate measurement of position. It therefore follows that for any sequential joint observable  $M$  on  $\mathbb{R}^2$  the inaccuracies satisfy the trade-off relations (45), (46) and (47) and, moreover, these relations constitute now the long-sought-for inaccuracy-disturbance trade-off relations.

Insofar as there are sequential measurement schemes in which these error and disturbance measures are finite, we have thus established the positive complement to the no-go statement (C): the associated sequential joint observable constitutes an approximate joint measurement, so that it is indeed possible to limit the disturbance of the momentum by allowing the position measurement to be only approximate.

The existence of sequential measurements of approximate position and momentum can be demonstrated by means of the “standard model” of an unsharp position measurement introduced by von Neumann (1932). In this model, the position of the object is measured by coupling it to the momentum  $\mathbf{P}_p$  of the probe system via  $U = e^{-(i/\hbar)\lambda\hat{Q}\otimes\hat{P}_p}$ , and using the position  $\mathbf{Q}_p$  of the probe as the readout observable. If  $\Psi_p$  is the initial probe state (assumed to be a bounded function), then the instrument of the measurement can be written in the form

$$\mathcal{I}(X)(\rho) = \int_X K_q \rho K_q^* dq, \quad (50)$$

with  $K_q$  denoting the multiplicative operator  $(K_q\psi)(x) = \sqrt{\lambda} \Psi_p(\lambda(q-x))\psi(x)$ . The approximate position realized by this measurement is  $\mathbf{Q}_\mu$ , where  $\mu$  is now the probability measure with distribution function  $\lambda|\Psi_p(-\lambda x)|^2$ .

Suppose now that one is carrying out first an approximate position measurement, with the instrument (50), and then a sharp momentum measurement. As shown by Davies (1970), this defines a unique sequential joint observable  $M$ , in fact, a covariant phase space observables with marginals

$$M_1(X) = \mathcal{I}^*(X)(\mathbf{P}(\mathbb{R})) = \mathbf{Q}_\mu(X), \quad M_2(Y) = \mathcal{I}^*(\mathbb{R})(\mathbf{P}(Y)) = \mathbf{P}_\nu(Y). \quad (51)$$

Here the distorted momentum is  $\mathbf{P}_\nu$ , where  $\nu$  is the probability measure with the distribution  $\frac{1}{\lambda}|\widehat{\Psi}_p(-\frac{p}{\lambda})|^2$ . It is obvious that  $\mu = \mu_T$  and  $\nu = \nu_T$ , where  $T = |\Psi^{(\lambda)}\rangle\langle\Psi^{(\lambda)}|$  with  $\Psi^{(\lambda)}(q) = \sqrt{\lambda}\Psi_p(\lambda q)$ . This makes it manifest that  $M$  obeys the uncertainty relations (45), (46) and (47), here in their double role as accuracy-accuracy and accuracy-disturbance trade-off relations.<sup>6</sup>

## VI. ILLUSTRATION: THE ARTHURS-KELLY MODEL

The best studied model of a joint measurement of position and momentum is that of Arthurs and Kelly (1965). In this model, a quantum object is coupled with two probe systems which are then independently measured to obtain information about the object’s position and momentum respectively. Arthurs and Kelly showed that this constitutes a simultaneous measurement of position and momentum in the sense that the distributions of the outputs reproduce the quantum expectation values of the object’s position and momentum. They also derived the uncertainty relation for the spreads of the output statistics corresponding to our Eq. (36). As shown in (Busch, 1982), the model also satisfies the more stringent condition of an approximate joint measurement, that the output statistics determine a covariant phase space observable whose marginals are smeared versions of position and momentum. This work also extended the model to a large class of probe input states (Arthurs and Kelly only considered Gaussian probe inputs),

<sup>6</sup> It is interesting to note that any covariant phase space observable can be formed in this way as a sequential joint measurement of  $\mathbf{Q}_\mu$  and  $\mathbf{P}$  (Carmeli *et al.*, 2006).

which made it possible to analyze the origin of the uncertainty relation for the measurement accuracies and identify the different relevant contributions to it. This will be described briefly below. For a detailed derivation of the induced observable and the state changes due to this measurement scheme, see (Busch, 1982) and (Busch *et al.*, 1997, Chapter 6). Further illuminating investigations of the Arthurs-Kelly model can be found, for instance, in (Stenholm, 1992) and (Raymer, 1994).

The Arthurs-Kelly model is based on the von Neumann model of an approximate position measurement introduced in Sec. V. The position  $\hat{Q}$  and momentum  $\hat{P}$  of the object are coupled with the position  $\hat{Q}_1$  and momentum  $\hat{P}_2$  of two probe systems, respectively, which serve as the readout observables. Neglecting the free evolutions of the three systems the combined time evolution is described by the measurement coupling

$$U = \exp\left(-\frac{i\lambda}{\hbar}\hat{Q} \otimes \hat{P}_1 \otimes I_2 + \frac{i\kappa}{\hbar}\hat{P} \otimes I_1 \otimes \hat{Q}_2\right). \quad (52)$$

If  $\psi$  is an arbitrary input (vector) state of the object, and  $\Psi_1, \Psi_2$  are the fixed initial states of the probes (given by suitable smooth functions, with zero expectations for the probes' positions and momenta), the probabilities for values of  $\hat{Q}_1$  and  $\hat{P}_2$  to lie in the intervals  $\lambda X$  and  $\kappa Y$ , respectively, determine a covariant phase space observable  $G^T$  of the form (32) via

$$\langle \psi | G^T(X \times Y) | \psi \rangle := \langle U\psi \otimes \Psi_1 \otimes \Psi_2 | I \otimes Q_1(\lambda X) \otimes P_2(\kappa Y) | U\psi \otimes \Psi_1 \otimes \Psi_2 \rangle. \quad (53)$$

The variances of the accuracy measures  $\mu, \nu$  associated with the marginals  $Q_\mu, P_\nu$  of  $G^T$  can readily be computed:

$$\begin{aligned} \Delta(\mu)^2 &= \frac{1}{\lambda^2} \Delta(\hat{Q}_1, \Psi_1)^2 + \frac{\kappa^2}{4} \Delta(\hat{Q}_2, \Psi_2)^2, \\ \Delta(\nu)^2 &= \frac{1}{\kappa^2} \Delta(\hat{P}_2, \Psi_2)^2 + \frac{\lambda^2}{4} \Delta(\hat{P}_1, \Psi_1)^2. \end{aligned} \quad (54)$$

If the two measurements did not disturb each other, only the first terms on the right hand sides would appear; the second terms are manifestations of the presence of the other probe and its coupling to the object. Since the observable defined in this measurement scheme is a covariant phase space observable, it follows immediately that the accuracy measures satisfy the trade-off relation (35),  $\Delta(\mu)\Delta(\nu) \geq \hbar/2$ . It is nevertheless instructive to verify this explicitly by evaluating the product of the above expressions:

$$\begin{aligned} \Delta(\mu)^2 \Delta(\nu)^2 &= \mathcal{Q} + \mathcal{D}, \\ \mathcal{Q} &:= \frac{1}{4} \Delta(\hat{Q}_1, \Psi_1)^2 \Delta(\hat{P}_1, \Psi_1)^2 + \frac{1}{4} \Delta(\hat{Q}_2, \Psi_2)^2 \Delta(\hat{P}_2, \Psi_2)^2 \geq \frac{\hbar^2}{8} \\ \mathcal{D} &:= \frac{1}{(\lambda\kappa)^2} \Delta(\hat{Q}_1, \Psi_1)^2 \Delta(\hat{P}_2, \Psi_2)^2 + \frac{(\lambda\kappa)^2}{16} \Delta(\hat{Q}_2, \Psi_2)^2 \Delta(\hat{P}_1, \Psi_1)^2 \geq \frac{\hbar^2}{16} \left(x + \frac{1}{x}\right) \geq \frac{\hbar^2}{8}, \\ \text{where } x &:= \frac{16}{(\lambda\kappa\hbar)^2} \Delta(\hat{Q}_1, \Psi_1)^2 \Delta(\hat{P}_2, \Psi_2)^2. \end{aligned} \quad (55)$$

Here we have repeatedly used the uncertainty relations for the probe systems,  $\Delta(\hat{Q}_k, \Psi_k)\Delta(\hat{P}_k, \Psi_k) \geq \hbar/2$ .

It is evident that there are two independent sources of inaccuracy in this joint measurement model. Indeed, each of the terms  $\mathcal{Q}$  and  $\mathcal{D}$  alone would suffice to guarantee an absolute positive lower bound for the inaccuracy product. The first term,  $\mathcal{Q}$ , is composed of two independent terms which reflect the *quantum nature* of the probe systems; there is no trace of a mutual influence of the two measurements being carried out simultaneously. This feature is in accordance with Bohr's argument concerning the possibilities of measurement, which he considered limited due to the quantum nature of parts of the measuring setup (the probe systems).

By contrast, the term  $\mathcal{D}$  reflects the mutual *disturbance* of the two measurements as it contains the coupling parameters and product combinations of variances associated with both probe systems. This feature of the mutual disturbance of measurements was frequently highlighted by Heisenberg in thought experiments aiming at joint or sequential determinations of the values of position and momentum.

A suitable modification of the measurement coupling  $U$  leads to a model that can be interpreted as a sequential determination of position and momentum. Consider the unitary operator, dependent on the additional real parameter  $\gamma$ ,

$$U^{(\gamma)} = \exp\left(-\frac{i\lambda}{\hbar}\hat{Q} \otimes \hat{P}_1 \otimes I_2 + \frac{i\kappa}{\hbar}\hat{P} \otimes I_1 \otimes \hat{Q}_2 - \frac{i\gamma\lambda\kappa}{2\hbar}I \otimes \hat{P}_1 \otimes \hat{Q}_2\right). \quad (56)$$

The Baker-Campbell-Hausdorff decomposition of this coupling yields

$$U^{(\gamma)} = \exp\left(-(\gamma+1)\frac{i}{2\hbar}\lambda\kappa I \otimes \hat{P}_1 \otimes \hat{Q}_2\right) \exp\left(-\frac{i}{\hbar}\lambda\hat{Q} \otimes \hat{P}_1 \otimes I_2\right) \exp\left(\frac{i}{\hbar}\kappa\hat{P} \otimes I_1 \otimes \hat{Q}_2\right). \quad (57)$$

It turns out that this coupling defines again a covariant phase space observable. The variances of the inaccuracy measures  $\mu_\gamma$ ,  $\nu_\gamma$  associated with the marginals are given as follows:

$$\begin{aligned} \Delta(\mu_\gamma)^2 &= \frac{1}{\lambda^2}\Delta(\hat{Q}_1, \Psi_1)^2 + (\gamma-1)^2\frac{\kappa^2}{4}\Delta(\hat{Q}_2, \Psi_2)^2, \\ \Delta(\nu_\gamma)^2 &= \frac{1}{\kappa^2}\Delta(\hat{P}_2, \Psi_2)^2 + (\gamma+1)^2\frac{\lambda^2}{4}\Delta(\hat{P}_1, \Psi_1)^2. \end{aligned} \quad (58)$$

These accuracies still satisfy the uncertainty relation (35), but this time the contributions corresponding to  $\mathcal{Q}$  and  $\mathcal{D}$  will both depend on the coupling parameters unless  $\kappa = 0$ . In particular, it does not help to make the coupling look like that of a sequential measurement, by putting  $\gamma = -1$ . In that case,  $\Delta(\nu_{-1})$  is the accuracy of an undisturbed momentum measurement, and  $\Delta(\mu_{-1})$  contains a term which reflects the disturbance of the subsequent position measurement through the momentum measurement. The disturbance of the position measurement accuracy is now given by  $\kappa\Delta(\hat{Q}_2, \Psi_2)$ , and together with the momentum inaccuracy it satisfies the uncertainty relation

$$\left[\frac{1}{\kappa^2}\Delta(\hat{P}_2, \Psi_2)^2\right] \left[\kappa^2\Delta(\hat{Q}_2, \Psi_2)^2\right] \geq \frac{\hbar^2}{4}. \quad (59)$$

## VII. ON EXPERIMENTAL IMPLEMENTATIONS AND TESTS OF THE UNCERTAINTY PRINCIPLE

“Turning now to the question of the empirical support [for the uncertainty principle], we unhesitatingly declare that rarely in the history of physics has there been a principle of such universal importance with so few credentials of experimental tests.” (Jammer, 1974, p. 81)

This assessment was written by the distinguished historian of physics Max Jammer at a time when studies of phase space observables based on positive operator measures were just beginning. He qualifies it with a survey of early proposed and actual tests of the preparation uncertainty relation, and he refers to some early model studies of joint measurements, the first of which being that by Arthurs and Kelly (1965).

Jammer’s verdict still holds true today. There are surprisingly few publications that address the question of experimental tests of the uncertainty principle. Some of these report confirmations of the uncertainty principle, while a few others predict or suggest violations. We will briefly comment on some of this work below.

### A. Tests of preparation uncertainty relations

The most commonly cited version of uncertainty relation is the preparation relation, usually in the familiar version in terms of standard deviations. Confirmations of this uncertainty relation have been reported by Shull (1969) for a single-slit diffraction experiment with neutrons, by Kaiser *et al.* (1983) and Klein *et al.* (1983) in neutron interferometric experiments, and more recently by Nairz *et al.* (2002) in a slit experiment for fullerene molecules.

In these slit diffraction and interferometric experiments, typical measures used for the width of the spatial wave function are the slit width and slit separation, respectively. The width of the associated momentum wave function is given in terms of the width at half height of the central peak. It must be noted that in the mathematical modeling of single slit diffraction, the standard deviation of the momentum distribution is infinite. Hence it is indeed necessary to use another, operationally significant measure of the width of that distribution. There does not seem to be a universally valid uncertainty relation involving width at half height (in short, half width), but the authors of these experiments make use of a Gaussian shape approximation of the central peak, which is in agreement with the data within the experimental accuracy. This allows them to relate the half widths to standard deviations and confirm the correct lower bound for the uncertainty product.

A model independent and thus more direct confirmation of the uncertainty principle can be obtained if the widths of the position and momentum distributions are measured in terms of the overall width defined in Eq. (4). It is likely that the data collected in these experiments contain enough information to determine these overall widths for different levels of total probability  $1 - \varepsilon_1$  and  $1 - \varepsilon_2$ . In the case of the neutron interference experiment, it was pointed out by Uffink (1985) that a more stringent relation is indeed at stake, namely, a trade-off relation, introduced by Uffink and

Hilgevoord (1985), between the overall width of the position distribution and the fine structure width (mean peak width) of the momentum distribution.

It should be noted that these experiments do not, strictly speaking, constitute *direct* tests of the uncertainty relations for position and momentum observables. While the position uncertainty, or the width of the position distribution, is determined as the width of the slit, the momentum distribution is inferred from the measured position distribution at a later time, namely when the particles hit the detection screen. This inference is based on the approximate far-field description of the wave function (Fraunhofer diffraction in optics), and is in accordance with the classical, geometric interpretation of momentum as mass times velocity. Thus, what is being tested is the uncertainty relation along with the free Schrödinger evolution and the Fourier-Plancherel connection between position and momentum.

An alternative interpretation can be given in the Heisenberg picture, noting that the operators  $\hat{Q}$ ,  $\hat{P}' := m\hat{Q}(t)/t$  are canonically conjugate, given the free evolution  $\hat{Q}(t) = \hat{Q} + \hat{P}t/m$ . Here  $m$  is the mass of the particle, and  $t$  is the time of passage of the wave packet from the slit to the detection screen. (If the distance between the slit and the detection screen is large compared to the longitudinal width of the wave packet, the time  $t$  is fairly well defined.) The width of the distribution of  $\hat{Q}$  is determined by the preparation (passage through the slit), and the distribution of  $\hat{P}'$  is measured directly.

## B. On implementations of joint and sequential measurements

To the best of our knowledge, and despite some claims to the contrary, there is presently no experimental realization of a joint measurement of position and momentum. Thus there can as yet be no question of an experimental test of the uncertainty relation for inaccuracies in joint measurements of these quantities. But there are reports on the successful experimental implementation of joint measurements of canonically conjugate quadrature components of quantum optical fields using multipoint homodyne detection.

There seem to be several communities in quantum optics and optical communication where these implementations were achieved independently. The experiment of Walker and Carrol (1986) is perhaps the first realization, with a theoretical analysis by Walker (1987) yielding the associated phase space observable. This seems to have been anticipated theoretically by Yuen and Shapiro (1980). See also (Lai and Haus, 1989) for a review. A more recent claim of a quantum optical realization of a joint measurement was made by Beck *et al.* (2001). It must be noted that in these works it is not easily established (in some cases for lack of sufficiently detailed information) whether the implementation criterion is merely that of reproducing the first moments of the two quadrature component statistics, or whether in fact the statistics of a joint observable have been measured.

By contrast, Freyberger *et al.* (1993), Freyberger and Schleich (1993) and Leonhardt and Paul (1993), and Leonhardt *et al.* (1995) showed that the eight port homodyne schemes for phase difference measurements carried out by Noh *et al.* (1991), Noh *et al.* (1992) yield statistics that approach the Q-function of the input state for a suitable macroscopic coherent state preparation of the local field mode. This is manifestly a realization of a joint observable. A simple analysis is given in (Busch *et al.*, 1997, Sec. VII.3.7.).

Turning to the question of position and momentum proper, the Arthurs-Kelly model is particularly well suited to elucidate the various aspects of the uncertainty principle for joint and, as we have seen, sequential joint measurements of approximate position and momentum. However, it is not clear whether and how an experimental realization of this scheme can be obtained. Apart from the quantum optical realizations of joint measurements of conjugate quadrature components, there are a few proposals of realistic schemes for position and momentum, e.g., (Royer, 1985), (Törmä *et al.*, 1995), and (Power *et al.*, 1997) mainly in the context of atom optics. In the latter two models the probe systems are electromagnetic field modes, and the readout probe observables are suitable phase-sensitive quantities. The measurement coupling differs from the Arthurs-Kelly coupling in accordance with the different choice of readout observables.

The experimental situation regarding the inaccuracy-versus-disturbance relation is far less well developed. This is probably because, as we have seen above, rigorous, operationally relevant formulations of such a relation had not been found until recently. Apart from some model considerations of the kind considered here in Sec. III there seems to be no experimental investigations of accuracy-disturbance trade-off relations.

## C. On some alleged violations of the uncertainty principle

Throughout the history of quantum mechanics, the joint measurement uncertainty relation has been the subject of repeated challenges. There are two lines of argument against it which start from logically contrary premises. The conclusion is, in either case, that only the preparation uncertainty relation is tenable (as a statistical relation) within quantum mechanics.

The first argument against the joint measurement relation was based on the claim that there is no provision for a notion of joint measurement within quantum mechanics. Based on a careful assessment of the attempts existing at the time, Ballentine (1970) concludes that a description of joint measurements of position and momentum in terms of joint probabilities could not be obtained without significant modifications or extensions of the existing theory. Here we have shown that the required modification was the introduction of positive operator measures and specifically phase space observables, which is entirely within the spirit of the traditional formulation of quantum mechanics; it amounts merely to a completion of the set of observables.

The second argument was based on the claim that joint measurements of position and momentum are in fact possible with arbitrary accuracy, and its authors, among them Karl Popper and Henry Margenau, attempted to demonstrate their claim by means of appropriate experimental schemes.

Popper (1934) conceived a joint measurement scheme that was based on measurements of entangled particle pairs. That this proposal was flawed and untenable was immediately noted by von Weizsäcker (1934). While Popper later accepted this criticism, he suggested (Popper, 1982, footnote on p. 15) that his example may nevertheless have inspired Einstein *et al.* (1935) to conceive the famous thought experiment. In fact, this experiment can be construed as a scheme for making a joint measurement of the position and momentum of a particle that is entangled with another particle in a particular state: *provided* that Einstein, Podolsky and Rosen's assumption of local realism is tenable, a measurement of the position of the latter particle allows one to infer the position of the first particle without disturbing that particle in any way. At the same time, one can then also measure the position of the first particle.

It would follow that the individual particle has definite values of position and momentum while quantum mechanics provides only an incomplete, statistical description. However, it is a well-known consequence of arguments such as the Kochen-Specker-Bell theorem (Bell, 1966; Kochen and Specker, 1967) and Bell's theorem (Bell, 1964) that such value assignments are in contradiction with quantum mechanics. Moreover, this contradiction has been experimentally confirmed in the case of Bell's inequalities, and these tests turned out in favor of quantum mechanics.

Another proposal of a joint determination of arbitrarily sharp values of the position and momentum of a quantum particle was made by Park and Margenau (1968) who considered the time of flight determination of velocity. As shown in a quantum mechanical analysis in (Busch and Lahti, 1984), this scheme is appropriately understood as a sequential measurement of first sharp position and then sharp momentum, and does therefore not constitute even an approximate joint measurement of position and momentum. But Park and Margenau are only interested in demonstrating that it is possible to ascribe arbitrarily sharp values of position and momentum to a single system at the same time.

An analogous situation arises in the slit experiment, where one could formally infer arbitrarily sharp values for the transversal momentum component from the bundle of geometric paths from any location in the slit to the detection point. This bundle is arbitrarily narrow if the separation between slit and detection screen is made sufficiently large. Thus the width of the spot on the detection screen and the width of the possible range of the inferred momentum value can be made small enough so that their product is well below the order of  $\hbar$ .

In both situations, the geometric reconstruction of a momentum value from the two position determinations at different times, which is guided by classical reasoning, constitutes an inference for the time between the two measurements and cannot be used to infer momentum distributions in the state before the measurement or to predict the outcomes of future measurements. Hence such values are purely formal and of no operational significance. One could be inclined to follow Heisenberg who noted in his 1929 Chicago lectures (Heisenberg, 1930, p. 25) that he regarded it as a matter of taste whether one considers such value assignments to past events as meaningful.

However, it has been shown, by an extension of the quantum mechanical language to incorporate propositions about past events, that hypothetical value assignments to past events lead to Kochen-Specker type contradictions. This result was obtained by Quadt (1988) in his diploma thesis written under P. Mittelstaedt's supervision at the University of Cologne in 1988; the argument is sketched in (Quadt, 1989).

Popper returned to the subject many years later (Popper, 1982, pp. 27-29) with a novel experimental proposal with which he aimed at testing (and challenging) the Copenhagen interpretation. In a subsequent experimental realization it is reported that the outcome seems to confirm Popper's prediction, thus amounting to an apparent violation of the preparation uncertainty relation.

In Popper's new experiment, EPR-correlated pairs of quantum particles are emitted from a source in opposite directions, and then each particle passes through a slit, a narrow one on one side, the one on the opposite side of wide opening. The particles are then recorded on a screen on each side. Popper predicts that independent diffraction patterns should build up on each side, according to the appropriate slit width; according to Popper, the Copenhagen interpretation should predict that the particle passing through the wider slit actually shows the same diffraction pattern as the other particle. In the extreme of no slit on one side, this would still be the case. Popper's interpretation of his experiment as a test of the Copenhagen interpretation was criticized soon afterwards, see, e.g., the exchange in (Collet and Loudon, 1987a,b; Popper, 1987) or (Sudbery, 1985).

The experimental realization of Popper's experiment by Kim and Shih (1999) shows, perhaps at first surprisingly, a behavior in line with Popper's prediction. Moreover, taking the width of the "ghost image" of the first, narrow slit

at the side of the second particle (confirmed by Pittman *et al.* (1995)) as a measure of the position uncertainty of the second particle, then this value together with the inferred width of the momentum distribution form a product smaller than allowed by the preparation uncertainty relation. Kim and Shih hasten to assert that this result does not constitute a *violation* of the uncertainty principle but is in agreement with quantum mechanics; still, the experiment has aroused some lively and controversial debate (e.g., (Bramon and Escribano, 2005; Peres, 2002; Sancho, 2002)). As pointed out by Short (2001), Kim and Shih overlook the fact that the two width parameters in question should be determined by the reduced quantum state of the particle and thus should, according to quantum mechanics, satisfy the uncertainty relation. Short gives an explanation of the experimental outcome in terms of the imperfect imaging process which leads to image blurring, showing that there is indeed no violation of the uncertainty relation.

Finally, it seems that papers with claims of actual or proposed experiments indicating violations of the uncertainty relation hardly ever pass the threshold of the refereeing process in major journals. They appear occasionally as contributions to conference proceedings dedicated to realistic (hidden variable) approaches to quantum mechanics.

## VIII. CONCLUSION

In this exposition we have elucidated the positive role of the uncertainty principle as a necessary and sufficient condition for the possibility of approximately localizing position and momentum. We have also noted that approximate position measurements can allow a control of the disturbance of the momentum. Uncertainty relations for position and momentum thus come in three variants: for the widths of probability distributions, for accuracies of joint measurements, and for the trade-off between the accuracy of a position measurement and the necessary momentum disturbance (and vice versa).

In his seminal paper of 1927, Heisenberg gave intuitive formulations of all three forms of uncertainty relations, but it was only the relation for state preparations that was made precise soon afterwards. It took several decades until the conceptual tools required for a rigorous formulation of the two measurement-related uncertainty relations had become available. Here we identified the following elements of such a rigorous formulation.

First, a theory of approximate joint measurements of position and momentum had to be developed; this possibility was opened up by the representation of observables as positive operator measures. Second, a criterion of what constitutes an approximate measurement of one observable by means of another must be based on operationally significant and experimentally relevant measures of inaccuracy or error. Here we discussed three candidate measures: standard error, a distance of observables, and error bars. For each of these, a universal Heisenberg uncertainty relation holds, showing that for any observable on phase space the marginal observables cannot both approximate position and momentum arbitrarily well.

The proofs of these uncertainty relations are first obtained for the distinguished class of covariant phase space observables, for which they follow mathematically from a form of uncertainty relation for state preparations. This formal connection between the preparation and joint measurement uncertainty relations is in accordance with a postulate formulated by Bohr (1928) in his famous Como lecture of 1927 which states that the possibilities of measurement should not exceed the possibilities of preparation. The uncertainty relation for a general approximate joint observable for position and momentum can then be deduced from that for some associated covariant joint observable.

Apart from the limitations on the accuracy of joint approximations of position and momentum, we have found Heisenberg uncertainty relations which quantify the necessary intrinsic unsharpness of two observables that are jointly measurable, *provided* they are to be approximations of position and momentum, respectively. Both limitations are consequences of the noncommutativity of position and momentum.

Finally, the idea of a measurement of (say) position disturbing the momentum has been made precise by recognizing that a sequential measurement of measuring first position and then momentum constitutes an instance of a joint measurement of some observables, of which the first marginal is an (approximate) position and the second a distorted momentum observable. The inaccuracy inherent in the second marginal gives a measure of the disturbance of momentum. The joint measurement uncertainty relations can in this context be interpreted as a trade-off between the accuracy of the first position measurement against the extent of the necessary disturbance of the momentum due to this measurement.

Last we have surveyed the current status of experimental implementations of joint measurements and the question of experimental tests of the uncertainty principle. While there do not seem to exist any confirmed violations of the uncertainty principle, there do exist several experimental tests of uncertainty relations which have shown agreement with quantum mechanics.

## Acknowledgement

Part of this work was carried out during mutual visits by Paul Busch at the University of Turku and by Pekka Lahti at Perimeter Institute. Hospitality and financial support by these host institutions are gratefully acknowledged. We are indebted to M. Leifer and R. Werner for helpful comments on an earlier draft version of this paper.

## References

- Appleby, D., 1998a, *Int. J. Theor. Phys.* **37**, 1491.
- Appleby, D., 1998b, *Int. J. Theor. Phys.* **37**, 2557.
- Arthurs, E., and J. Kelly, 1965, *Bell System Tech. J.* **44**, 725.
- Ballentine, L., 1970, *Rev. Mod. Phys.* **42**, 358.
- Beck, M., C. Dorner, and I. Walmsley, 2001, *Phys. Rev. Lett.* **87**, 253601/1.
- Bell, J., 1964, *Physics* **1**, 195.
- Bell, J., 1966, *Rev. Mod. Phys.* **38**, 447.
- Bohr, N., 1928, *Nature* **121**, 580.
- Bramon, A., and R. Escribano, 2005, [quant-ph/0501134](https://arxiv.org/abs/quant-ph/0501134).
- Busch, P., 1982, *Unbestimmtheitsrelation und simultane Messungen in der Quantentheorie*, Ph.D. thesis, University of Cologne, english translation: Indeterminacy relations and simultaneous measurements in quantum theory, *Int. J. Theor. Phys.* **24** (1985) 63–92.
- Busch, P., M. Grabowski, and P. Lahti, 1997, *Operational Quantum Physics* (Springer-Verlag, Berlin), second corrected printing.
- Busch, P., T. Heinonen, and P. Lahti, 2004, *Phys. Lett. A* **320**, 261.
- Busch, P., and P. Lahti, 1984, *Phys. Rev. D* **29**, 1634.
- Busch, P., P. Lahti, and P. Mittelstaedt, 1996, *The Quantum Theory of Measurement* (Springer-Verlag, Berlin), second revised edition.
- Busch, P., and D. Pearson, 2006, Universal joint-measurement uncertainty relation for error bars, unpublished – available at [quant-ph/0612XXX](https://arxiv.org/abs/quant-ph/0612XXX).
- Busch, P., T. Schonbek, and F. Schroeck, Jr., 1987, *J. Math. Phys.* **28**, 2866.
- Carmeli, C., T. Heinonen, and A. Toigo, 2004, *J. Math. Phys.* **45**, 2526.
- Carmeli, C., T. Heinonen, and A. Toigo, 2006, Intrinsic unsharpness and approximate repeatability of quantum measurements, unpublished – available at [quant-ph/0611286](https://arxiv.org/abs/quant-ph/0611286).
- Cassinelli, G., E. De Vito, and A. Toigo, 2003, *J. Math. Phys.* **44**, 4768.
- Collet, M., and R. Loudon, 1987a, *Nature* **326**, 671.
- Collet, M., and R. Loudon, 1987b, *Nature* **328**, 675.
- Cowling, M., and J. Price, 1984, *SIAM J. Math. Anal.* **15**, 151.
- Davies, E., 1970, *J. Funct. Anal.* **6**, 318.
- Davies, E., 1976, *Quantum Theory of Open Systems* (Academic Press, London).
- Dym, H., and H. McKean, 1972, *Fourier Series and Integrals* (Academic Press, New York).
- Einstein, A., B. Podolsky, and N. Rosen, 1935, *Phys. Rev.* **47**, 777.
- Folland, G., and A. Sitaram, 1997, *J. Fourier Anal. Appl.* **3**, 207.
- Freyberger, M., and W. Schleich, 1993, *Phys. Rev. A* **47**, R30.
- Freyberger, M., K. Vogel, and W. Schleich, 1993, *Phys. Lett. A* **176**, 41.
- Hall, M., 2004, *Phys. Rev. A* **69**, 052113/1.
- Heisenberg, W., 1927, *Z. Phys.* **43**, 172.
- Heisenberg, W., 1930, *The Physical Principles of the Quantum Theory* (University of Chicago Press, Chicago).
- Heisenberg, W., 1977, in *The Uncertainty Principle and Foundations of Quantum Mechanics. A Fifty Years' Survey*, edited by W. Price and S. Chissick (J. Wiley & Sons, London), pp. 3–6.
- Holevo, A., 1979, *Rep. Math. Phys.* **16**, 385.
- Jammer, M., 1974, *The Philosophy of Quantum Mechanics* (John Wiley & Sons, New York).
- Kaiser, H., S. Werner, and E. George, 1983, *Phys. Rev. Lett.* **50**, 560.
- Kim, Y.-H., and Y. Shih, 1999, *Found. Phys.* **29**, 1849.
- Kiukas, J., P. Lahti, and K. Ylinen, 2006a, *J. Math. Anal. Appl.* **319**, 783.
- Kiukas, J., P. Lahti, and K. Ylinen, 2006b, *J. Math. Phys.* **47**, 072104.
- Klein, A., G. Opat, and W. Hamilton, 1983, *Phys. Rev. Lett.* **50**, 563.
- Kochen, S., and E. Specker, 1967, *J. Math. Mech.* **17**, 59.
- Lahti, P., 1986, *Rep. Math. Phys.* **23**, 289.
- Lai, Y., and H. Haus, 1989, *Quantum Opt.* **1**, 99.
- Landau, H., and H. Pollak, 1961, *Bell System Tech. J.* **40**, 65.
- Lenard, A., 1972, *J. Functional Analysis* **10**, 410.
- Leonhardt, U., B. Böhmer, and H. Paul, 1995, *Opt. Commun.* **119**, 296.
- Leonhardt, U., and H. Paul, 1993, *Phys. Rev. A* **47**, R2460.
- Nairz, O., M. Arndt, and A. Zeilinger, 2002, *Phys. Rev. A* **65**, 032109/1.
- von Neumann, J., 1932, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin), english translation: *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton, 1955.
- Noh, J., A. Fougères, and L. Mandel, 1991, *Phys. Rev. Lett.* **67**, 1426.
- Noh, J., A. Fougères, and L. Mandel, 1992, *Phys. Rev. A* **45**, 424.
- Ozawa, M., 2004, *Ann. Physics* **311**, 350.
- Park, J., and H. Margenau, 1968, *Int. J. Theor. Phys.* **1**, 211.
- Peres, A., 2002, *Stud. Hist. Philos. Modern Phys.* **33**, 23.

- Pittman, T., Y. Shih, D. Strekalov, and A. Sergienko, 1995, *Phys. Rev. A* **52**, R3429.
- Popper, K., 1934, *Naturwiss.* **22**, 807.
- Popper, K., 1982, *Quantum Theory and the Schism in Physics* (Unwin Hyman, London).
- Popper, K., 1987, *Nature* **328**, 675.
- Power, W., S. Tan, and M. Wilkens, 1997, *J. Mod. Opt.* **44**, 2591.
- Quadt, R., 1988, *Über die Nicht-Objektivierbarkeit quantenmechanischer Ereignisse in der Vergangenheit*, Diploma thesis, University of Cologne.
- Quadt, R., 1989, *Found. Phys.* **19**, 1027.
- Raymer, M., 1994, *Am. J. Phys.* **62**, 986.
- Royer, A., 1985, *Phys. Rev. Lett.* **55**, 2745.
- Sancho, P., 2002, *Found. Phys.* **32**, 789.
- Scully, M., B.-G. Englert, and H. Walther, 1991, *Nature* **351**, 111.
- Short, A., 2001, *Found. Phys. Lett.* **14**, 275.
- Shull, C., 1969, *Phys. Rev.* **179**, 752.
- Stenholm, S., 1992, *Ann. Physics* **218**, 233.
- Sudbery, A., 1985, *Phil. Sci.* **52**, 470.
- Törmä, P., S. Stenholm, and I. Jex, 1995, *Phys. Rev. A* **52**, 4812.
- Uffink, J., 1985, *Phys. Lett.* **108A**, 59.
- Uffink, J., and J. Hilgevoord, 1985, *Found. Phys.* **15**, 925.
- Walker, N., 1987, *J. Mod. Optics* **34**, 15.
- Walker, N., and J. Carrol, 1986, *Optical and Quantum Electronics* **18**, 355.
- von Weizsäcker, K., 1934, *Naturwiss.* **22**, 808.
- Werner, R., 1984, *J. Math. Phys.* **25**, 1404.
- Werner, R., 2004, in *Quantum information, Statistics, Probability*, edited by O. Hirota (Paramus, NJ:Rinton), pp. 153–171, also available: [quant-ph/0405184](http://quant-ph/0405184).
- Ylinen, K., 1989, *J. Math. Anal. Appl.* **137**, 185.
- Yuen, H., and J. Shapiro, 1980, *IEEE Trans. Inform. Theory* **26**, 78.