Disturbance, conservation laws and the uncertainty principle

1992 J. Phys. A: Math. Gen. 25 4887

(http://iopscience.iop.org/0305-4470/25/18/021)

J. Phys. A: Math. Gen. 25 (1992) 4887-4901. Printed in the UK

Disturbance, conservation laws and the uncertainty principle

Hans Martens and Willem M de Muynck

Department of Theoretical Physics, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands

Received 18 December 1991, in final form 27 April 1992

Abstract. The interpretation of the uncertainty principle in terms of a measurement of a *single* observable disturbing other observables, originating in Heisenberg's 1927 paper, is shown to be derivable from an uncertainty principle for *joint* measurements of incompatible observables. The latter also limits measurements of a single observable in the presence of conservation laws.

1. Introduction

In Heisenberg's classic 1927 paper [1], the *uncertainty principle* (UP) is introduced with the aid of the so-called γ-microscope. The set-up is sketched in figure 1. In order to determine the horizontal component q of the position of a microscopic particle, it is illuminated with light of wavelength λ. The scattered light is collected through a lens with aperture ε onto a photographic plate. The resolution of the microscope, in other words its inaccuracy as a position meter, is given by

$$\delta_q \sim \lambda/\sin\varepsilon. \tag{1}$$

On the other hand, the light imparts a certain amount of momentum to the electron, due to the Compton effect. The horizontal component p of the final momentum is only known up to [2]

$$D_p \sim \sin\varepsilon/\lambda \tag{2}$$

($\hbar = 1$). We might say that the indeterminateness of the recoil causes a *disturbance* of the particle's momentum. The product of the momentum disturbance and the resolution satisfies

$$D_p\delta_q \sim 1. \tag{3}$$

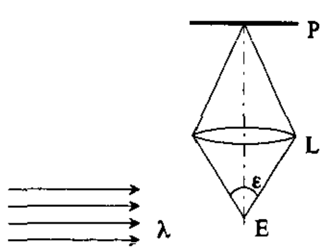

Figure 1. Heisenberg's γ-microscope *Gedanken* experiment. Light with wavelength λ falls on a particle E. The scattered light is collected by a lens L with aperture ε onto a photographic plate P.

0305-4470/92/184887+15$04.50

Heisenberg uses this reasoning as a heuristic argument for the more formal relation that bears his name, namely

$$\langle\Delta^2\hat{P}\rangle\langle\Delta^2\hat{Q}\rangle \geq \tfrac{1}{4} \tag{4}$$

(operators are careted; $\langle\Delta^2\hat{P}\rangle$ denotes $\langle(\hat{P}-\langle\hat{P}\rangle)^2\rangle$). But on closer inspection the connection between (4) and (3) turns out to be not quite that obvious. On the one hand, according to Born's statistical interpretation, relation (4) limits statistical scatter or spread [3], and both variances derive directly from the object state $|\psi\rangle$. In (3), on the other hand, δ_q quantifies a property of the measuring instrument, rather than that it is connected to the state of a microscopic particle. Moreover, it characterizes the measurement's inaccuracy: the microscope is *not* assumed to be a perfectly accurate position meter. It does not leave the object in a state with sharp q, and when the object is prepared in a state with sharp q, the measurement does not give that result with certainty. Therefore the γ-microscope is not a measurement of the first kind, is indeed not at all a measurement of a self-adjoint operator along the lines of the von Neumann axiomatization. But, on the contrary, inasfar as measurement is involved in (4), which is only implicitly, it is measurement as axiomatized by von Neumann.

Physicists have, notwithstanding the aforementioned problems, applied uncertainty relations to situations outside the domain assigned to scatter-type relations like (4) by Born's interpretation. But the use of uncertainty relations that are, like Heisenberg's γ-microscope relation (3), only informally established, has its disadvantages. Due to lack of mathematical exactness, it is often not quite clear precisely what limit is implied by a given type of UP. Thus discussion results, as for example in the case of the supposed quantum limits to path measurements [4]. In particular the disturbance version of the UP, the version that is illustrated by the γ-microscope, has been questioned [5]. Because of the central place of the UP in quantum mechanics, clarity about its meaning is important, however.

In the present paper we therefore aim to show how uncertainty relations like (3) can be derived formally. First we give a formalism, generalizing von Neumann's, in which imperfect measurements like the γ-microscope can be discussed. Within this formalism an uncertainty principle has been derived [6, 7], that limits the accuracy achievable in joint measurements of incompatible observables. Relations similar to (3) can, as will be shown, in their turn be derived from the latter type of UP, thus confirming Heisenberg's intuition in a formal way.

It has long been known that conservation laws restrict the possibilities of measurement [8–11]. This has been seen as an extra limitation on quantum measurements, in addition to the limits given by the UP. As a matter of fact, however, limits induced by conservation laws can be derived from the UP. We end the paper with formal illustrations of both the γ-microscope and the conservation law limit.

2. Non-ideal measurements

In realistic situations, measurements are not described by self-adjoint operators, or *projection-valued measures* (PVMs), but rather by *positive operator-valued measures* (POVMs) [11, 12]. The POVM notion forms an extension of the measurement description of von Neumann's axioms. For a discrete outcome set K, a POVM $\mathfrak{m} = \{\hat{M}_k\}$ $(k \in K)$ is a set of operators satisfying

$$\hat{M}_k \geq \hat{0} \qquad \sum_k \hat{M}_k = \hat{1}. \tag{5}$$

In general $[\hat{M}_k, \hat{M}_l]_- \neq \hat{0}$ if $k \neq l$. Moreover, as the operator $\hat{M}_k$ need not satisfy $\hat{M}_k = \hat{M}_k^2$, it is not necessarily a projector. If the object state corresponds to the density operator $\hat{\rho}$, the probability that a realization of $\mathfrak{m}$ yields outcome k is given by $\mathrm{Prob}_{\mathfrak{m}}(k) := \mathrm{Tr}(\hat{\rho}\hat{M}_k)$.

If we measure a POVM $\mathfrak{m} = \{\hat{M}_k\}$, this may be seen as a non-ideal measurement of a PVM $\mathfrak{f} = \{\hat{F}_l\}$ if there exists a matrix (λ_{kl}) such that

$$\hat{M}_k = \sum_l \lambda_{kl}\hat{F}_l \qquad \lambda_{kl} \geq 0 \qquad \sum_k \lambda_{kl} = 1. \tag{6}$$

Thus the matrix (λ_{kl}) is a stochastic matrix. Definition (6) implies that the $\mathfrak{m}$-probabilities $\mathrm{Prob}_{\mathfrak{m}}(k)$ are completely determined by the $\mathfrak{f}$-probabilities $\mathrm{Prob}_{\mathfrak{f}}(l) = \mathrm{Tr}(\hat{\rho}\hat{F}_l)$, whereas the reverse is in general untrue. More specifically, the distribution $\mathrm{Prob}_{\mathfrak{m}}(k)$ is a 'smeared' version of the $\mathfrak{f}$ distribution $\mathrm{Prob}_{\mathfrak{f}}(l)$. Even if the object state is such that the $\mathfrak{f}$ distribution is sharp, a measurement of $\mathfrak{m}$ will in general not give one result with certainty. The matrix (λ_{kl}) is a characteristic of the $\mathfrak{m}$-device. The device will be least non-ideal if λ_{kl} is the Kronecker-δ. If on the other hand λ_{kl} depends only on k $(\lambda_{kl} = \tilde{\lambda}_k)$, the non-ideality is maximal. Then $\mathrm{Prob}_{\mathfrak{m}}(k)$ is independent of $\mathfrak{f}$, is in fact independent of $\hat{\rho}$. Note finally that non-ideality (6) may be used to define a partial ordering on the set of POVMs [13]. We shall in the following use the shorthand notation $\mathfrak{f} \to \mathfrak{m}$ for (6). The non-ideality notion was introduced by Davies [11] and by Prugovečki and co-workers [14], and studied more systematically by the authors in [13] (see also [15, 16] and other references in [13]).

For the continuous case, an analogue of (6) can be given. Consider a PVM $\mathfrak{e} = \{\hat{E}(x)\,\mathrm{d}x\}$ on $\mathbb{R}$. Then

$$\hat{O}(y) = \int_{\mathbb{R}} \lambda_x(y)\hat{E}(x)\,\mathrm{d}x \qquad \lambda_x(y) \geq 0 \qquad \int_{\mathbb{R}} \lambda_x(y)\,\mathrm{d}y = 1 \tag{7}$$

defines a POVM $\mathfrak{o}$. It is related to $\mathfrak{e}$ analogous to the way $\mathfrak{f}$ is related to $\mathfrak{m}$ in (6). In other words, $\mathfrak{e} \to \mathfrak{o}$. In the special case that $\lambda_x(y) = \bar{\lambda}(y-x)$, (7) becomes a convolution [11, 13], namely

$$\hat{O}(y) = \int_{\mathbb{R}} \bar{\lambda}(y-x)\hat{E}(x)\,\mathrm{d}x. \tag{8}$$

3. Uncertainty principle for joint measurements

In this section we shall study the joint measurement of two PVMs on the finite dimensional Hilbert space $\mathbb{C}^n$. All characteristic elements as regards incompatibility are present in such a space. Moreover, physically relevant results for infinite dimensional cases may be obtained through the $n \to \infty$ limit.

A bivariate POVM $\mathfrak{r} = \{\hat{R}_{mj}\}$ represents a joint non-ideal measurement of two PVMs $\mathfrak{e} = \{\hat{E}_k\}$ and $\mathfrak{f} = \{\hat{F}_l\}$ if its marginals represent non-ideal measurements of $\mathfrak{e}$ and $\mathfrak{f}$, respectively. In other words, it represents a joint non-ideal measurement if there are matrices (λ_{mk}) and (μ_{jl}) such that

$$\begin{aligned} \hat{R}_m^{(1)} &= \sum_j \hat{R}_{mj} = \sum_k \lambda_{mk}\hat{E}_k \qquad \lambda_{mk} \geq 0 \qquad \sum_m \lambda_{mk} = 1 \\ \hat{R}_j^{(2)} &= \sum_m \hat{R}_{mj} = \sum_l \mu_{jl}\hat{F}_l \qquad \mu_{jl} \geq 0 \qquad \sum_j \mu_{jl} = 1. \end{aligned} \tag{9}$$

The POVMs $\mathfrak{r}^{(1)}=\{\hat{R}^{(1)}_m\}$ and $\mathfrak{r}^{(2)}=\{\hat{R}^{(2)}_j\}$ denote $\mathfrak{r}$'s marginals. Equations (9) may thus be equivalently written in the form $\mathfrak{e}\to\mathfrak{r}^{(1)}\wedge\mathfrak{f}\to\mathfrak{r}^{(2)}$. It has been shown [6] that such an $\mathfrak{r}$ exists for arbitrary $\mathfrak{e},\mathfrak{f}$.

An uncertainty principle for joint measurements should imply that the matrices (λ_{mk}) and (μ_{jl}) in (9) cannot both approach the Kronecker-δ arbitrarily closely, if $\mathfrak{e}$ and $\mathfrak{f}$ are incompatible. In order to derive such an uncertainty principle, we must quantify the *amount* of non-ideality present in a given measurement, i.e. how much a given matrix (λ_{kl}) differs from the Kronecker-δ. One possible non-ideality measure, derived from the *conditional* entropy of information theory, is

$$J_{\mathfrak{f}\to\mathfrak{m}}:=-\sum_k\sum_l\lambda_{kl}P_l\log\left[\frac{\lambda_{kl}P_l}{\sum_m\lambda_{km}P_m}\right]\qquad P_l=\frac{1}{n}\mathrm{Tr}(\hat{F}_l)\tag{10}$$

$[\mathfrak{f},\mathfrak{m}$ as in (6)]. J is non-negative. In accord with the interpretation of non-ideality and the intended meaning of J, it is zero if $\lambda_{kl}=\delta_{kl}$; if $\lambda_{kl}=\tilde{\lambda}_k$, J takes its maximal value. The better the measurement, the lower J. It can be shown [6, 7, 17] that

$$J_{\mathfrak{e}\to\mathfrak{r}^{(1)}}+J_{\mathfrak{f}\to\mathfrak{r}^{(2)}}\geqslant\sum_i\mathrm{Tr}(\hat{G}_i)c_i\tag{11}$$

where

$$c_i=-\frac{1}{n}\log\left(\max_{k,l}\|\hat{E}_k\hat{F}_l\hat{G}_i\|^2\right)$$

where $\{\hat{G}_i\}$ is a PVM that commutes with both $\mathfrak{e}$ and $\mathfrak{f}$. (This PVM is introduced to resolve cases where $\mathfrak{e}$ and $\mathfrak{f}$ have eigenspaces in common.) The right-hand side of (11) is zero if and only if $\mathfrak{e}$ and $\mathfrak{f}$ are compatible (for an optimal choice of $\{\hat{G}_i\}$). Thus (11) presents a non-trivial lower bound for the inaccuracy achievable in joint measurements of two incompatible observables.

On infinite-dimensional spaces, no result as general as (11) is known as yet. But for position–momentum, an uncertainty relation can be derived for joint measurements with convolution-type marginals, as in (8). We speak of *covariant* non-ideal measurements [6, 12, 15]. For non-ideality of that type, an analogue of (10) is†

$$J_{\mathfrak{e}\to\mathfrak{o}}:=-\int_{\mathbb{R}}\bar{\lambda}(y)\log[\bar{\lambda}(y)].\tag{12}$$

Then, combining a result of Holevo's for covariant joint non-ideal measurements ([12], theorem IV.8.1) with an entropic variant [19] of the uncertainty relation (4), we can derive the following inaccuracy relation:

$$J_{\mathfrak{p}\to\mathfrak{t}^{(1)}}+J_{\mathfrak{q}\to\mathfrak{t}^{(2)}}\geqslant 1+\log(\pi).\tag{13}$$

Here $\mathfrak{q}$ and $\mathfrak{p}$ denote the position and momentum PVMs, respectively. The POVM $\mathfrak{t}$ realizes a joint non-ideal measurement of $\mathfrak{p}$ and $\mathfrak{q}$ along the lines of (8) and (9), $\mathfrak{t}^{(1)}$ and $\mathfrak{t}^{(2)}$ being marginals of $\mathfrak{t}$. Note that the use of entropic quantities in both (13) and (11) is not crucial. What is needed is simply some expression of the amount of non-ideality in a given function $\bar{\lambda}(y)$ or matrix λ_{kl}. Hence for non-ideality measures other than J, relations analogous to (13) and (11) should also be derivable. Indeed

† The discrete analogue of a convolution involves a square matrix in which the rows are cyclic permutations of each other [6]. Then it follows that [18]

$$J_{\mathfrak{f}\to\mathfrak{m}}=-\sum_k\lambda_{kl}\log(\lambda_{kl})$$

for an arbitrary l, independent of the distribution (P_l). This latter relation corresponds to (12).

non-entropic versions of (13) have been derived by Ali and Prugovečki [15], and by others (e.g. [12, 16], cf also [11]). Such relations, although *quantitatively* different, express the same *qualitative* principle: that the inaccuracy with which incompatible observables are measurable jointly, is limited.

4. Operations

As we saw, Heisenberg's γ-microscope involves momentum *disturbance* as well as position *inaccuracy.* Because in the thought experiment the particle's momentum is not actually measured, disturbance is different from measurement inaccuracy. It is the momentum of the outgoing particle that is disturbed with respect to the particle's initial state. Therefore 'disturbance' involves the state of the outgoing particle, i.e. the state of the object after the measurement. In the previous section, we described measurements with POVMs. This description did not take the output state into account, because that was not necessary for a characterization of non-ideality. Now we will first extend the measurement description to include the output state. Then, in the following section, we shall see that the disturbance problem can be phrased in terms of consecutive measurements. Consecutive measurements can be seen as joint measurements. Hence, even though disturbance is different from inaccuracy, the formalism for joint measurements outlined above, can be applied to the disturbance issue. In this way we are able to confirm Heisenberg's reasoning formally.

First consider a simple model of the measurement. The object, in state $\hat{\rho}_{\mathcal{O}}$, is coupled to an apparatus in state $\hat{\rho}_{\mathcal{A}}$. Then an interaction, corresponding to the unitary operator $\hat{U}$, entangles these states to $\hat{\rho}_{\mathcal{O}+\mathcal{A},\text{final}}$. Schematically,

$$\hat{\rho}_{\mathcal{O}} \to \hat{\rho}_{\mathcal{O}} \otimes \hat{\rho}_{\mathcal{A}} \to \hat{\rho}_{\mathcal{O}+\mathcal{A},\text{final}} = \hat{U}\hat{\rho}_{\mathcal{O}} \otimes \hat{\rho}_{\mathcal{A}}\hat{U}^{\dagger}. \tag{14}$$

Next some apparatus observable, the pointer observable, is read out. We take this to be some PVM (or POVM), say $\{\hat{C}_k\}$. Partial tracing over $\mathcal{A}$ shows the object output state, conditional on outcome k, to become

$$\hat{\rho}_{\mathcal{O}+\mathcal{A}} \to \hat{\hat{\rho}}_{\mathcal{O},k} = \mathrm{Tr}_{\mathcal{A}}[\hat{\rho}_{\mathcal{O}+\mathcal{A},\text{final}}\hat{C}_k]. \tag{15}$$

The probability of outcome k is thus

$$\mathrm{Tr}[\hat{\rho}_{\mathcal{O}+\mathcal{A},\text{final}}\hat{C}_k] = \mathrm{Tr}[\hat{\hat{\rho}}_{\mathcal{O},k}] = \mathrm{Tr}[\hat{\rho}_{\mathcal{O}}\hat{M}_k] \tag{16}$$

where

$$\hat{M}_k = \mathrm{Tr}_{\mathcal{A}}[\hat{\rho}_{\mathcal{A}}\hat{U}^{\dagger}\hat{C}_k\hat{U}].$$

Accordingly, the measurement's POVM $\{\hat{M}_k\}$ on the $\mathcal{O}$ Hilbert space follows naturally from the model [11, 20]. But the full description of this measurement is by the transition

$$\hat{\rho}_{\mathcal{O}} \to \hat{\mu}_k[\hat{\rho}_{\mathcal{O}}] := \hat{\hat{\rho}}_{\mathcal{O},k} = \mathrm{Tr}_{\mathcal{A}}[\hat{U}\hat{\rho}_{\mathcal{O}} \otimes \hat{\rho}_{\mathcal{A}}\hat{U}^{\dagger}\hat{C}_k]. \tag{17}$$

The mapping (17) is in fact an *operation valued measure* [11] (OVM). An *operation* is a linear mapping $\hat{\phi}$ from the set of trace-class operators into the set of trace-class operators, satisfying [20]

$$\forall_{\hat{\rho}|\hat{\rho}\geqslant\hat{0}}\, \hat{\phi}[\hat{\rho}] \geqslant \hat{0} \qquad \forall_{\hat{\rho}|\hat{\rho}\geqslant\hat{0}}\, \mathrm{Tr}(\hat{\phi}[\hat{\rho}]) \leqslant \mathrm{Tr}(\hat{\rho}). \tag{18}$$

We call an operation $\hat{\phi}$ *non-selective* if it satisfies, in addition to (18),

$$\forall_{\hat{\rho}|\hat{\rho}\geqslant\hat{0}}\, \mathrm{Tr}(\hat{\phi}[\hat{\rho}]) = \mathrm{Tr}(\hat{\rho}). \tag{19}$$

Now, an OVM $\mathfrak{M}=\{\hat{\mu}_k\}$ ($k\in K$) is defined as a set of operations $\hat{\mu}_k$ such that $\hat{\mu}_K := \Sigma_{k\in K}\,\hat{\mu}_k$ is a non-selective operation. The mapping (17) can easily be seen to satisfy this requirement. As in (16), the norm $\mathrm{Tr}(\hat{\mu}_k[\hat{\rho}])$ corresponds to the probability of obtaining result k. Hence the measurement's POVM $\mathfrak{m}=\{\hat{M}_k\}$ is uniquely defined by

$$\mathrm{Tr}(\hat{\rho}\hat{M}_k)=\mathrm{Tr}(\hat{\mu}_k[\hat{\rho}]) \Leftrightarrow \hat{M}_k=\hat{\mu}_k^{\dagger}[\hat{1}] \tag{20}$$

(the adjoint being defined by $\mathrm{Tr}(\hat{\phi}[\hat{\tau}]\hat{A})=\mathrm{Tr}(\hat{\tau}\hat{\phi}^{\dagger}[\hat{A}])$, $\hat{\tau}$ trace-class, $\hat{A}$ bounded [20]). Note that to a given OVM there corresponds only one POVM, but that there may be many OVMs for one given POVM. The von Neumann–Lüders [21] first kind measurement for a PVM $\mathfrak{e}=\{\hat{E}_k\}$,

$$\hat{\rho}\to\hat{E}_k\hat{\rho}\hat{E}_k =: \hat{\varepsilon}_k[\hat{\rho}] \tag{21}$$

is a special kind of OVM, satisfying $\hat{\varepsilon}_k\{\hat{\varepsilon}_l[\hat{\rho}]\}=\delta_{kl}\hat{\varepsilon}_k[\hat{\rho}]$ for all $\hat{\rho}$.

5. Disturbance

Perhaps the first concrete meaning for 'disturbance' to come to mind, is an increase of the statistical spread in the output state for observables that are incompatible with the one measured. But, as Kraus [5] has shown, not all OVMs have this effect. Imagine, Kraus argues, a device that measures spin. It first rotates the spins 90° around the x-axis, and then measures $\mathfrak{s}_y$ as in (21). Here $\mathfrak{s}_y$ denotes the PVM corresponding to the self-adjoint operator $\hat{\sigma}_y$. Further denote $\hat{\sigma}_y$ ($\hat{\sigma}_z$) eigenstates by $|\to\rangle$ and $|\leftarrow\rangle$ ($|\uparrow\rangle$ and $|\downarrow\rangle$). Then Kraus's measurement transforms the spin state according to

$$\hat{\rho}\to\hat{\rho}_k \qquad k=\text{up, down} \tag{22}$$

$$\hat{\rho}_{\text{up}}=|\to\rangle\langle\uparrow|\hat{\rho}|\uparrow\rangle\langle\to| \qquad \hat{\rho}_{\text{down}}=|\leftarrow\rangle\langle\downarrow|\hat{\rho}|\downarrow\rangle\langle\leftarrow|.$$

The apparatus measures $\mathfrak{s}_z$ accurately: if the input state is a $\hat{\sigma}_z$-eigenstate, the corresponding measurement result ensues with certainty. Nevertheless $\hat{\sigma}_y$ is sharp in the output state, conditional on the measurement outcome. The scatter in $\hat{\sigma}_y$ is certainly not increased.

But note that all information about the $\hat{\sigma}_y$ distribution in the *initial* state has been wiped out. Whatever measurement we next perform on the final state, we learn nothing about the original $\hat{\sigma}_y$ distribution. Analogously, we see that in the γ-microscope experiment a measurement of the momentum of the outgoing particle is not a good measurement of the initial momentum, due to the indeterminate recoil. In the following we shall therefore interpret 'disturbance' in terms of this loss of information. It will be shown that a measurement on the final state of another, incompatible, measurement is always less informative about the initial state than a measurement of the corresponding observable on the initial state itself. Against this form of 'disturbance' objections like Kraus's do not hold, as opposed to an interpretation in terms of increased scatter.

Let us now investigate this disturbance notion in a more formal way. Suppose we perform, on an object in state $\hat{\rho}$, a measurement corresponding to the OVM $\mathfrak{M}=\{\hat{\mu}_k\}$ (with POVM $\mathfrak{m}=\{\hat{M}_k\}$). When the object has left the $\mathfrak{M}$ apparatus, we measure the POVM $\mathfrak{n}=\{\hat{N}_l\}$ on it. Then the $\mathfrak{n}$ distribution is given by

$$\mathrm{Tr}(\hat{\mu}_K[\hat{\rho}]\hat{N}_l)=\mathrm{Tr}(\hat{\rho}\hat{\mu}_K^{\dagger}[\hat{N}_l]) =: \mathrm{Prob}_{\mathfrak{n}\circ\mathfrak{M}}(l) \tag{23}$$

as follows from $\mathfrak{M}$'s definition in the previous section. Thus, the second measurement can be seen as a measurement of the POVM $\mathfrak{n}\circ\mathfrak{M} := \hat{\mu}_K^{\dagger}[\mathfrak{n}]=\{\hat{\mu}_K^{\dagger}[\hat{N}_l]\}$ on the initial object state $\hat{\rho}$. If we perform this measurement in order to find out something about the PVM $\mathfrak{f}$ in the initial state, we should therefore choose $\mathfrak{n}$ such that $\mathfrak{f} \to \mathfrak{n} \circ \mathfrak{M}$, rather than $\mathfrak{f} \to \mathfrak{n}$. Only in special cases, e.g. if

$$\mathfrak{f} \to \mathfrak{f} \circ \mathfrak{M} = \hat{\mu}_K^\dagger[\mathfrak{f}] \tag{24}$$

can the two requirements be expected to coincide. In case (24) is valid, the $\mathfrak{f}$ distribution is smeared. But, as Kraus's reasoning shows, there is no fundamental reason why OVMs in general should satisfy a requirement like (24). Indeed (22) does not. In the same vein we see, for example, in a non-destructive measurement of photon number [22] a disturbance of phase that is not directly connected to an increase in phase spread, or to a smearing of the phase distribution.

Accordingly, if we want to determine the initial $\mathfrak{f}$, when $\mathfrak{M}$ has already done its work, we will aim at minimizing the non-ideality in $\mathfrak{f} \to \mathfrak{n} \circ \mathfrak{M} = \hat{\mu}_K^\dagger[\mathfrak{n}]$. Using J as a criterion, we get the definition

$$J_{\hat{\mu}_K}^{\delta\mathfrak{f}} := \inf(J_{\mathfrak{f} \to \hat{\mu}_K^\dagger[\mathfrak{n}]}) \tag{25}$$

the infimum being taken over the class of POVMs $\mathfrak{n}$ satisfying $\mathfrak{f} \to \hat{\mu}_K^\dagger[\mathfrak{n}]$. The quantity $J_{\hat{\mu}_K}^{\delta\mathfrak{f}}$ represents the least inaccuracy with which the initial $\mathfrak{f}$ can be measured after the $\mathfrak{M}$ measurement has been performed. It is a measure for the amount of $\mathfrak{f}$ disturbance in the $\mathfrak{M}$-realization. Take the spin measurement discussed above as an example. We see from (22) that there the trivial, uninformative POVM $\{\hat{1}\}$ is the only choice for $\mathfrak{n}$ that satisfies $\hat{\sigma}_y \to \hat{\mu}_K^\dagger[\mathfrak{n}]$. Hence $J_{\hat{\mu}_K}^{\delta\hat{\sigma}_y}$ takes the largest possible value, log(2), despite the fact that the final $\mathfrak{S}_y$ distribution is sharp conditional on the measurement outcome.

Note that (25) is defined for the non-selective operation $\hat{\mu}_K$, and can be applied to any non-selective operation, whether it is connected to an OVM or not. In particular, a unitary transformation $\hat{\Xi}[\hat{\rho}] = \hat{U}\hat{\rho}\hat{U}^\dagger$ is a non-selective operation. It has zero disturbance: $J_{\hat{\Xi}}^{\delta\mathfrak{a}} = 0$ for any POVM $\mathfrak{a}$. The reason for this is that we could, at least in principle, measure the POVM

$$\bar{\mathfrak{a}} := \hat{\Xi}^\dagger[\mathfrak{a}] = \hat{U}\mathfrak{a}\hat{U}^\dagger \tag{26}$$

on the final state, to obtain unimpaired information on $\mathfrak{a}$ in the initial state. Thus, we see that not every change of $\hat{\rho}$ is a disturbance in the sense of (25). We might refer to changes that may be undone analogous to (26), as *distortion*, to distinguish them from disturbance, which cannot be undone in this way. Thus, whereas $\hat{\Xi}$ from equation (26) distorts without disturbing, we may view (24) as a case where $\mathfrak{f}$ is only disturbed, and not distorted. Generally, however, a measurement transformation will involve both distortion and disturbance. The spin example (22) is a case in point. The purpose of the somewhat involved formulation of (25) is therefore precisely to separate disturbance from distortion, to give the amount of disturbance in a given transformation regardless of the possible presence of distortion.

Finally, we shall relate disturbance to inaccuracy in a relation like (3). Let $\mathfrak{M}$ measure some PVM $\mathfrak{e}$ non-ideally (i.e. $\mathfrak{e} \to \mathfrak{m}$). The results of $\mathfrak{M}$ and a subsequently performed $\mathfrak{n}$-measurement give a bivariate distribution $\mathrm{Prob}_{\mathfrak{m}\,\&\,\mathfrak{n}\circ\mathfrak{M}}(k, l)$,

$$\mathrm{Prob}_{\mathfrak{m}\,\&\,\mathfrak{n}\circ\mathfrak{M}}(k, l) = \mathrm{Tr}(\hat{\mu}_k[\hat{\rho}]\hat{N}_l) = \mathrm{Tr}(\hat{\rho}\hat{\mu}_k^\dagger[\hat{N}_l]). \tag{27}$$

The marginals of this bivariate POVM $\{\hat{\mu}_k^\dagger[\hat{N}_l]\}$ are $\mathfrak{m}$ and $\mathfrak{n} \circ \mathfrak{M} = \hat{\mu}_K^\dagger[\mathfrak{n}]$. Thus, if $\mathfrak{n}$ satisfies $\mathfrak{f} \to \hat{\mu}_K^\dagger[\mathfrak{n}]$, the two consecutive measurements together form a joint non-ideal measurement of $\mathfrak{e}$ and $\mathfrak{f}$. Since for all such $\mathfrak{n}$ the inaccuracy relation (11) must hold, it follows that the disturbance (25) must satisfy

$$J_{\mathfrak{e}\to\mathfrak{m}}+J^{\delta\mathfrak{f}}_{\hat{\mu}_K}\geqslant\sum_i \mathrm{Tr}(\hat{G}_i)c_i. \tag{28}$$

From (28) we see that, the lower the inaccuracy $J_{\mathfrak{e}\to\mathfrak{m}}$ is, the stronger the lower bound on the disturbance $J^{\delta\mathfrak{f}}_{\hat{\mu}_K}$. Thus the full validity of Heisenberg's disturbance idea is demonstrated, contrary to Kraus's skepticism.

Application of the joint measurement UP (13) for position-momentum gives, in a completely analogous way, a connection between the position inaccuracy and the momentum disturbance. Again, the entropic nature of (28), like that of (11) and (13) is not fundamental. If we start from other versions of these latter relations, different forms of (28) result. These would, although quantitatively different, express the same general principle as (28).

6. Example

We next consider a measurement model to illustrate the disturbance idea, and its application in a disturbance-inaccuracy relation like (28). In contrast to other treatments [4, 23] we take a rather more abstract route. That has the advantage that the assumptions which go into the reasoning can be clearly distinguished. In this way we can see which aspects of the model are general, and which are not. We shall first present the reasoning leading to the disturbance-inaccuracy relation for this particular case, briefly highlighting the assumptions made on the way. After this, their meaning is discussed more fully, as are other aspects of the model.

As in section 4, we describe the measurement process by means of a pointer system $\mathscr{A}$ and a read-out observable $\mathfrak{c}=\{\hat{C}(q)\,\mathrm{d}q\}$ on $\mathscr{H}_{\mathscr{A}}$. If the unitary operator $\hat{U}$ describes the interaction between system $\mathscr{A}$ and the object $\mathscr{O}$, then the device's OVM $\mathfrak{A}=\{\hat{a}(q)\,\mathrm{d}q\}$ and POVM $\mathfrak{a}=\{\hat{A}(q)\,\mathrm{d}q\}$ are given by

$$\begin{aligned}\hat{a}(q)[\hat{\rho}_{\mathscr{O}}]&=\mathrm{Tr}_{\mathscr{A}}[\hat{U}\hat{\rho}_{\mathscr{O}}\otimes\hat{\rho}_{\mathscr{A}}\hat{U}^{\dagger}\hat{C}(q)]\\ \hat{A}(q)=\hat{a}^{\dagger}(q)[\hat{1}]&=\mathrm{Tr}_{\mathscr{A}}[\hat{\rho}_{\mathscr{A}}\hat{U}^{\dagger}\hat{C}(q)\hat{U}].\end{aligned} \tag{29}$$

In the γ-microscope the system $\mathscr{A}$ directly interacting with the particle $\mathscr{O}$ is the light field. It is customary to take this field monochromatic, so as to have a well defined initial momentum. This can be formulated as a condition of invariance of the initial $\mathscr{A}$-state $\hat{\rho}_{\mathscr{A}}$ under translations, yielding as our first assumption that $\hat{\rho}_{\mathscr{A}}$ satisfies

$$\hat{S}_{\mathscr{A}}(x)\hat{\rho}_{\mathscr{A}}\hat{S}^{\dagger}_{\mathscr{A}}(x)=\hat{\rho}_{\mathscr{A}} \tag{30}$$

for all† x, the unitary operator $\hat{S}_{\mathscr{A}}(x)$ corresponding to a position shift of system $\mathscr{A}$ by an amount x. In the γ-microscope, the read-out of the position measurement is performed by the lens and the photographic plate. Accordingly, the read-out observable is here taken to have a continuous outcome set, $\mathfrak{c}=\{\hat{C}(q)\,\mathrm{d}q\}$, which is covariant under a simultaneous position shift of both object and measuring apparatus. This is expressed by our second assumption to the effect that the read-out observable satisfies the equality

$$\hat{S}^{\dagger}_{\mathscr{O}}(x)\otimes\hat{S}^{\dagger}_{\mathscr{A}}(x)\hat{C}(q)\hat{S}_{\mathscr{O}}(x)\otimes\hat{S}_{\mathscr{A}}(x)=\hat{C}(q-x) \tag{31}$$

† By restricting the assumption of translation invariance to a finite interval it is possible to deal with the problem that states satisfying (29) for all x are not normalizable. For the sake of simplicity we shall not consider this here.

$\hat{S}_{\mathcal{O}}(x)$ corresponding to a position shift of the object $\mathcal{O}$. By (31) it is expressed that the effect of a position shift x of both object and measuring instrument on the probability distribution of the read-out observable can be described by a translation of the coordinate system. Because $\hat{c}$ only operates on $\mathcal{H}_{\mathcal{A}}$, condition (31) is equivalent to covariance with respect to $\hat{S}_{\mathcal{A}}(x)$ alone.

For simplicity's sake we assume that $\mathcal{O}$ has only one degree of freedom (or, equivalently, that the meter is insensitive to the other ones). Further we take an interaction in which total momentum is conserved.

Finally we assume the interaction to satisfy the equality

$$\hat{U}\hat{Q}_{\mathcal{O}}\otimes\hat{1}_{\mathcal{A}}\hat{U}^{\dagger}=\hat{\hat{Q}}_{\mathcal{O}}\otimes\hat{1}_{\mathcal{A}} \tag{32}$$

for some operator $\hat{\hat{Q}}_{\mathcal{O}}$ on $\mathcal{H}_{\mathcal{O}}$. This condition can easily be seen to be equivalent to the equality

$$\mathrm{Tr}_{\mathcal{O}}\,\hat{\rho}_{\mathcal{O}}\hat{Q}_{\mathcal{O}}=\mathrm{Tr}_{\mathcal{O}}\,\hat{\hat{Q}}_{\mathcal{O}}\,\mathrm{Tr}_{\mathcal{A}}\,\hat{U}\hat{\rho}_{\mathcal{O}}\otimes\hat{\rho}_{\mathcal{A}}\hat{U}^{\dagger}$$

which means that the measured observable (i.e. position $\hat{Q}_{\mathcal{O}}$), is itself not disturbed in the sense of section 5. In other words, the device does not hamper a subsequently performed position measurement. This assumption is implicit in other models [4, 23]. It implies a generalized form of the so-called measurement of the first kind, satisfying

$$|q\rangle_{\mathcal{O}}\otimes|\psi\rangle_{\mathcal{A}}\rightarrow\hat{U}|q\rangle_{\mathcal{O}}\otimes|\psi\rangle_{\mathcal{A}}=|\bar{q}\rangle_{\mathcal{O}}\otimes|\phi(q)\rangle_{\mathcal{A}} \tag{33}$$

where $\hat{\rho}_{\mathcal{A}}=|\psi\rangle_{\mathcal{A}}\langle\psi|$, and $|\bar{q}\rangle_{\mathcal{O}}$ denotes an eigenstate of $\hat{\hat{Q}}_{\mathcal{O}}$ at eigenvalue q. It is clear from this requirement that the states $|\bar{q}_1\rangle_{\mathcal{O}}$ and $|\bar{q}_2\rangle_{\mathcal{O}}$ must be orthogonal if $|q_1\rangle_{\mathcal{O}}$ and $|q_2\rangle_{\mathcal{O}}$ are.

In order to satisfy (30), $|\psi\rangle_{\mathcal{A}}$ must be invariant with respect to $\hat{S}_{\mathcal{A}}(\Delta q)$ up to a phase factor $\exp(\mathrm{i}\varphi\Delta q)$. This, together with conservation of total momentum in the interaction between systems $\mathcal{O}$ and $\mathcal{A}$, makes it possible to obtain the final state in (33), up to a phase factor, by means of a position shift from the state at the fixed point $q=0$, i.e.,

$$\begin{aligned}|\bar{q}\rangle_{\mathcal{O}}\otimes|\phi(q)\rangle_{\mathcal{A}}&=\exp(\mathrm{i}\varphi\Delta q)\hat{S}_{\mathcal{O}}(\Delta q)\otimes\hat{S}_{\mathcal{A}}(\Delta q)|\overline{q-\Delta q}\rangle_{\mathcal{O}}\otimes|\phi(q-\Delta q)\rangle_{\mathcal{A}}\\&=\exp(\mathrm{i}\varphi q)\hat{S}_{\mathcal{O}}(q)\otimes\hat{S}_{\mathcal{A}}(q)|\bar{0}\rangle_{\mathcal{O}}\otimes|\phi(0)\rangle_{\mathcal{A}}.\end{aligned} \tag{34}$$

Finally, from the covariance requirement (31) it follows that our model satisfies the useful equality

$${}_{\mathcal{A}}\langle\phi(0)|\hat{C}(q-q')|\phi(0)\rangle_{\mathcal{A}}={}_{\mathcal{A}}\langle\phi(0)|\hat{S}^{\dagger}_{\mathcal{A}}(q')\hat{C}(q)\hat{S}_{\mathcal{A}}(q')|\phi(0)\rangle_{\mathcal{A}}. \tag{35}$$

This completes our model. We now want to study its properties as regards position measurement inaccuracy and momentum disturbance. We first consider position measurement inaccuracy. One possibility for performing an inaccurate position measurement would be by measuring the position $\hat{Q}_{\mathcal{A}}$ of $\mathcal{A}$ in the final state. Then (33) would yield

$$\mathrm{Prob}(q)=\int \mathrm{d}q'|_{\mathcal{A}}\langle q|\phi(q')\rangle_{\mathcal{A}}|^2|\psi_{\mathcal{O}}(q')|^2$$

$\psi_{\mathcal{O}}(q')$ being the initial object wavefunction. It is clear from this expression that the inaccuracy of the position measurement is caused by the spreading of measurement results q for given q'.

More generally, it can be shown that an observable $\mathfrak{c}$ as defined above represents a non-ideal measurement of position. Thus, using (29) to define the POVM that is actually measured, it follows straightforwardly from (32)-(35) that

$$\mathrm{Tr}[\hat{\rho}_{\mathcal{O}}\hat{A}(q)] = \int_{\mathbb{R}} \mathrm{d}q' \langle q'|\hat{\rho}_{\mathcal{O}}|q'\rangle_{\mathcal{A}}\langle\phi(0)|\hat{C}(q-q')|\phi(0)\rangle_{\mathcal{A}}. \tag{36}$$

Comparing (36) to (8) we see that $\mathfrak{a}$ is a non-ideal measurement of $\mathfrak{q}_{\mathcal{O}}$ with ${}_{\mathcal{A}}\langle\phi(0)|\hat{C}(q-q')|\phi(0)\rangle_{\mathcal{A}}$ as the 'smearing' function $\bar{\lambda}$. From (35) and (36) we conclude that the measurement inaccuracy is directly connected to how good the POVM $\{\hat{C}(q)\,\mathrm{d}q\}$ can distinguish the states $\hat{S}_{\mathcal{A}}(q')|\phi(0)\rangle_{\mathcal{A}}$ [23]. Take as a measure of distinguishability the function (proposed by Wootters [24])

$$\begin{aligned} U(\Delta q) &= \int_{\mathbb{R}} [{}_{\mathcal{A}}\langle\phi(0)|\hat{C}(q-\Delta q)|\phi(0)\rangle_{\mathcal{A}}\,{}_{\mathcal{A}}\langle\phi(0)|\hat{C}(q)|\phi(0)\rangle_{\mathcal{A}}]^{1/2}\,\mathrm{d}q \\ &= \int_{\mathbb{R}} [{}_{\mathcal{A}}\langle\phi(0)|\hat{S}^{\dagger}_{\mathcal{A}}(\Delta q)\hat{C}(q)S_{\mathcal{A}}(\Delta q)|\phi(0)\rangle_{\mathcal{A}} \\ &\quad \times {}_{\mathcal{A}}\langle\phi(0)|\hat{C}(q)|\phi(0)\rangle_{\mathcal{A}}]^{1/2}\,\mathrm{d}q. \end{aligned} \tag{37}$$

Whereas $U(0)=1$, with increasing Δq the function U is expected to fall off. The faster this goes, the better $\mathfrak{c}$ distinguishes the pointer states. Of course $\mathfrak{c}$ should be chosen such that optimal distinguishability is achieved, i.e. such that U is as low as possible for $\Delta q \neq 0$. For given $|\phi(0)\rangle_{\mathcal{A}}$ the distinguishability function U cannot be decreased indefinitely, however: it has been shown by Hilgevoord and Uffink [25] (cf [26]) that U is bounded by

$$U(\Delta q) \geqslant |{}_{\mathcal{A}}\langle\phi(0)|\hat{S}_{\mathcal{A}}(\Delta q)|\phi(0)\rangle_{\mathcal{A}}| \tag{38}$$

thus implying also a lower bound to the accuracy of the position measurement described by (36).

Next we determine the momentum disturbance. For this the following result, which can be derived from (34) for the OVM $\mathfrak{A}$ defined in (29), is helpful:

$$\begin{aligned} \mathrm{Tr}[\hat{\alpha}(\mathbb{R})[\hat{\rho}_{\mathcal{O}}]\exp(-\mathrm{i}x\hat{P}_{\mathcal{O}})] &= \mathrm{Tr}[\hat{\alpha}(\mathbb{R})[\hat{\rho}_{\mathcal{O}}]\hat{S}_{\mathcal{O}}(x)] \\ &= u(x)\,\mathrm{Tr}[\hat{\rho}_{\mathcal{O}}\exp(-\mathrm{i}x\hat{P}_{\mathcal{O}})] \end{aligned} \tag{39}$$

$$u(x) = \exp(-\mathrm{i}\varphi x)\,{}_{\mathcal{A}}\langle\phi(0)|\hat{S}^{\dagger}_{\mathcal{A}}(x)|\phi(0)\rangle_{\mathcal{A}}.$$

Since $\hat{\alpha}(\mathbb{R})[\hat{\rho}_{\mathcal{O}}]$ is just the final object state, this implies that in the course of the measurement the amplitude of the Fourier transform of the object momentum distribution is multiplied by a factor $u(x)$, $|u| \leqslant 1$. From this it directly follows by taking the Fourier transform that the final momentum distribution is a convolution of the initial one with a positive distribution function. Evidently, by the measuring process discussed here momentum is disturbed in the sense of being 'smeared', the disturbance being larger as $u(x)$ deviates more from 1.

The absolute value $|u|$ of the factor in (39) may be recognized as the right-hand side of equation (38). This latter inequality shows that good distinguishability of the pointer states requires a low value for $|u|$. Substituting that low value into (39), we see that the Fourier transform of the momentum distribution is reduced significantly, thus implying a large momentum disturbance. Hence, momentum disturbance and position inaccuracy are complementary in the sense that by improving one the other is made worse. This is in complete agreement with the conclusions of section 5 (although there we used different measures to characterize the amounts of disturbance and inaccuracy in the general result from the ones we employ in this example).

This measurement model has, like the γ-microscope itself, a number of special properties. Firstly, we return to our third assumption, (32). As follows from (32) and (33), the PVM of $\hat{\hat{Q}}_{\mathcal{O}}$ is the observable we must measure on the final $\mathcal{O}$-state in order to obtain information on $\mathcal{O}$-position in the initial state. More generally, it can be shown that the model satisfies

$$\mathrm{Tr}[\hat{\alpha}(\mathbb{R})[\hat{\rho}_{\mathcal{O}}]f(\hat{\hat{Q}}_{\mathcal{O}})] = \mathrm{Tr}[\hat{\rho}_{\mathcal{O}}f(\hat{Q}_{\mathcal{O}})] \tag{40}$$

for arbitrary functions f. Therefore (32) implies that the position observable $\mathfrak{q}_{\mathcal{O}}$ is not disturbed in the sense of (25): $J^{\delta \mathfrak{q}_0} = 0$ [cf (26)]. Secondly, this model has the special property that momentum information remains in the momentum distribution, as can be seen from (39). In other words, a momentum measurement on the final object state is a non-ideal momentum measurement for the initial object state [cf equation (24)]:

$$\mathfrak{p}_{\mathcal{O}} \to \hat{\alpha}^{\dagger}(\mathbb{R})[\mathfrak{p}_{\mathcal{O}}]. \tag{41}$$

Contrary to the complementarity between disturbance and inaccuracy as expressed by (28), both of these special properties are not necessarily representative of measurements in general. We shall now investigate to which properties of the device they are connected. First consider position non-disturbance. Our device was intended to measure position, i.e. $\mathfrak{q}_{\mathcal{O}} \to \mathfrak{a}$. Since this implies that the operators $\hat{A}(q)$ defined in (29) must commute with the position operator $\hat{Q}_{\mathcal{O}}$, this requirement is equivalent to invariance of the POVM with respect to momentum shifts [13]:

$$\hat{A}(q) = \hat{T}^{\dagger}_{\mathcal{O}}(p)\hat{A}(q)\hat{T}_{\mathcal{O}}(p) \tag{42}$$

the unitary operator $\hat{T}_{\mathcal{O}}(p)$ effecting a momentum shift by an amount p. From the right-hand side of (42) we have

$$\begin{aligned}
\mathrm{Tr}[\hat{A}(q)\hat{T}_{\mathcal{O}}(p)\hat{\rho}_{\mathcal{O}}\hat{T}^{\dagger}_{\mathcal{O}}(p)] &= \mathrm{Tr}[\hat{U}[\hat{T}_{\mathcal{O}}(p)\hat{\rho}_{\mathcal{O}}\hat{T}^{\dagger}_{\mathcal{O}}(p)]\otimes\rho_{\mathcal{A}}\hat{U}^{\dagger}\hat{C}(q)] \\
&= \mathrm{Tr}[\hat{U}\hat{T}_{\mathcal{O}}(p)[\hat{\rho}_{\mathcal{O}}\otimes\hat{\rho}_{\mathcal{A}}]\hat{T}^{\dagger}_{\mathcal{O}}(p)\hat{U}^{\dagger}\hat{C}(q)] \\
&= \mathrm{Tr}[\hat{\hat{T}}(p)\hat{U}\hat{\rho}_{\mathcal{O}}\otimes\hat{\rho}_{\mathcal{A}}\hat{U}^{\dagger}\hat{\hat{T}}^{\dagger}(p)\hat{C}(q)] \\
&= \mathrm{Tr}[\hat{U}\hat{\rho}_{\mathcal{O}}\otimes\hat{\rho}_{\mathcal{A}}\hat{U}^{\dagger}\hat{\hat{T}}^{\dagger}(p)\hat{C}(q)\hat{\hat{T}}(p)]
\end{aligned} \tag{43}$$

with $\hat{\hat{T}}(p)$ defined by

$$\hat{\hat{T}}(p) = \hat{U}\hat{T}_{\mathcal{O}}(p)\hat{U}^{\dagger} = \exp(\mathrm{i}p\hat{\hat{Q}}) \tag{44}$$

and $\hat{\hat{Q}} = \hat{U}\hat{Q}_{\mathcal{O}}\otimes\hat{1}_{\mathcal{A}}\hat{U}^{\dagger}$. Note that in general $\hat{\hat{Q}}$ is *not* operating exclusively on the object variables. Hence, comparing with (32) we see that $\hat{\hat{Q}} \neq \hat{\hat{Q}}_{\mathcal{O}}\otimes\hat{1}_{\mathcal{A}}$ in general. From (42) and (43) it follows that a sufficient condition for $q_{\mathcal{O}} \to \mathfrak{a}$ is:

$$\hat{\hat{T}}^{\dagger}(p)\hat{C}(q)\hat{\hat{T}}(p) = \hat{C}(q) \Leftrightarrow [\hat{C}(q), \hat{\hat{Q}}]_{-} = \hat{0}. \tag{45}$$

If $\hat{\hat{Q}}$ is an operator on $\mathcal{H}_{\mathcal{O}}$ alone, as is the case in (32), equation (45) is automatically fulfilled, because $\mathfrak{c}$ operates on $\mathcal{H}_{\mathcal{A}}$. In conclusion, we have shown that position *non-disturbance* (32) is a *sufficient* condition for the meter to *measure* position. Position non-disturbance is, however, not a necessary condition for $\mathfrak{c}$ to represent a non-ideal position measurement. Thus, if the states $|\phi(q)\rangle_{\mathcal{A}}$ in (33) are orthogonal for different values of q, the states $|\bar{q}\rangle_{\mathcal{O}}$ may be allowed to be non-orthogonal. Nevertheless, any $\mathfrak{c}$ that is diagonal in the $|\phi(q)\rangle_{\mathcal{A}}$ representation describes a non-ideal position measurement.

We now consider the second special property, equation (41). Using the three general properties, namely, the invariance property (30), momentum conservation and the covariance property (31), we see that:

$$
\begin{aligned}
\hat{\alpha}(q)[\hat{S}_{\mathcal{O}}(x)\hat{\rho}_{\mathcal{O}}\hat{S}^{\dagger}_{\mathcal{O}}(x)]
&= \mathrm{Tr}_{\mathscr{A}}[\hat{U}[\hat{S}_{\mathcal{O}}(x)\hat{\rho}_{\mathcal{O}}\hat{S}^{\dagger}_{\mathcal{O}}(x)]\otimes\hat{\rho}_{\mathscr{A}}\hat{U}^{\dagger}\hat{C}(q)] \\
&= \mathrm{Tr}_{\mathscr{A}}[\hat{U}[\hat{S}_{\mathcal{O}}(x)\hat{\rho}_{\mathcal{O}}\hat{S}^{\dagger}_{\mathcal{O}}(x)]\otimes[\hat{S}_{\mathscr{A}}(x)\hat{\rho}_{\mathscr{A}}\hat{S}^{\dagger}_{\mathscr{A}}(x)]\hat{U}^{\dagger}\hat{C}(q)] \\
&= \mathrm{Tr}_{\mathscr{A}}[\hat{S}_{\mathcal{O}}(x)\otimes\hat{S}_{\mathscr{A}}(x)\hat{U}\hat{\rho}_{\mathcal{O}}\otimes\hat{\rho}_{\mathscr{A}}\hat{U}^{\dagger}\hat{S}^{\dagger}_{\mathcal{O}}(x)\otimes\hat{S}^{\dagger}_{\mathscr{A}}(x)\hat{C}(q)] \\
&= \hat{S}_{\mathcal{O}}(x)\,\mathrm{Tr}_{\mathscr{A}}[\hat{U}\hat{\rho}_{\mathcal{O}}\otimes\hat{\rho}_{\mathscr{A}}\hat{U}^{\dagger}\hat{S}^{\dagger}_{\mathcal{O}}(x)\otimes\hat{S}^{\dagger}_{\mathscr{A}}(x)\hat{C}(q)\hat{S}_{\mathcal{O}}(x)\otimes\hat{S}_{\mathscr{A}}(x)]\hat{S}^{\dagger}_{\mathcal{O}}(x) \\
&= \hat{S}_{\mathcal{O}}(x)\,\mathrm{Tr}_{\mathscr{A}}[\hat{U}\hat{\rho}_{\mathcal{O}}\otimes\hat{\rho}_{\mathscr{A}}\hat{U}^{\dagger}\hat{C}(q-x)]\hat{S}^{\dagger}_{\mathcal{O}}(x) \\
&= \hat{S}_{\mathcal{O}}(x)\hat{\alpha}(q-x)[\hat{\rho}_{\mathcal{O}}]\hat{S}^{\dagger}_{\mathcal{O}}(x).
\end{aligned}
\tag{46}
$$

Therefore the OVM is *covariant* with respect to $\mathcal{O}$ position shifts,

$$
\hat{\alpha}(q)[\hat{S}_{\mathcal{O}}(x)\hat{\rho}_{\mathcal{O}}\hat{S}^{\dagger}_{\mathcal{O}}(x)] = \hat{S}_{\mathcal{O}}(x)\hat{\alpha}(q-x)[\hat{\rho}_{\mathcal{O}}]\hat{S}^{\dagger}_{\mathcal{O}}(x) \tag{47}
$$

implying

$$
\hat{S}^{\dagger}_{\mathcal{O}}(x)\hat{A}(q)\hat{S}_{\mathcal{O}}(x) = \hat{A}(q-x) \tag{48}
$$

for the POVM $\mathfrak{a}$. Physically, covariance of the OVM means that the response to a shifted input state is also shifted, as is the output state. It has the consequence that $\hat{\alpha}(q)$'s adjoint satisfies the equality

$$
\begin{aligned}
\int_{\mathbb{R}} \hat{S}^{\dagger}_{\mathcal{O}}(x)\hat{\alpha}^{\dagger}(q)[\hat{P}_{\mathcal{O}}]\hat{S}_{\mathcal{O}}(x)\,\mathrm{d}q &\equiv \hat{S}^{\dagger}_{\mathcal{O}}(x)\hat{\alpha}^{\dagger}(\mathbb{R})[\hat{P}_{\mathcal{O}}]\hat{S}_{\mathcal{O}}(x) \\
&= \hat{\alpha}^{\dagger}(\mathbb{R})[\hat{S}^{\dagger}_{\mathcal{O}}(x)\hat{P}_{\mathcal{O}}\hat{S}_{\mathcal{O}}(x)] = \hat{\alpha}^{\dagger}(\mathbb{R})[\hat{P}_{\mathcal{O}}]
\end{aligned}
\tag{49}
$$

which means that $[\hat{\alpha}^{\dagger}(\mathbb{R})[\hat{P}_{\mathcal{O}}], \hat{P}_{\mathcal{O}}]_{-} = \hat{0}$. Thus, if $\mathscr{A}$ is insensitive to $\mathcal{O}$'s other degrees of freedom, OVM covariance (47) implies equation (41). Hence in the γ-microscope $\mathfrak{p}$ disturbance goes, as a consequence of the device's covariance, hand in hand with an increased $\hat{P}$-scatter.

7. Conservation laws

Consider a measuring apparatus $\mathscr{A}$, with OVM $\mathfrak{A}$ and POVM $\mathfrak{a}$, which measures position: $\mathfrak{q}_{\mathcal{O}} \to \mathfrak{a}$. $\mathscr{A}$ interacts with an object $\mathcal{O}$ such that total momentum $\hat{P}_{\mathrm{tot}} = \hat{P}_{\mathscr{A}} + \hat{P}_{\mathcal{O}}$ is conserved. In principle it is possible to use a second meter to measure $\hat{P}_{\mathrm{tot}}$ after $\mathscr{A}$ has performed its measurement. Let $\mathfrak{p}_{\mathrm{tot}} := \{\hat{E}_{p_{\mathrm{tot}}}(p)\,dp\}$ denote the PVM corresponding to $\hat{P}_{\mathrm{tot}}$ (and analogously for the PVMs of $\hat{P}_{\mathscr{A}}$ and $\hat{P}_{\mathcal{O}}$), and assume apparatus and object to be initially in states $\hat{\rho}_{\mathscr{A}}$ and $\hat{\rho}_{\mathcal{O}}$, respectively. For the $\mathfrak{p}_{\mathrm{tot}}$ outcome distribution we can then write

$$
\begin{aligned}
\mathrm{Prob}_{\mathfrak{p}_{\mathrm{tot}}}(p) &= \mathrm{Tr}[\hat{\rho}_{\mathscr{A}+\mathcal{O},\mathrm{final}}\hat{E}_{p_{\mathrm{tot}}}(p)] = \mathrm{Tr}[\hat{\rho}_{\mathscr{A}}\otimes\hat{\rho}_{\mathcal{O}}\hat{E}_{p_{\mathrm{tot}}}(p)] \\
&= \int_{\mathbb{R}} \mathrm{Tr}[\hat{\rho}_{\mathscr{A}}\hat{E}_{p_{\mathscr{A}}}(p-p')]\,\mathrm{Tr}[\hat{\rho}_{\mathcal{O}}\hat{E}_{p_{\mathcal{O}}}(p')]\,\mathrm{d}p' \\
&=: \int_{\mathbb{R}} \lambda(p-p')\,\mathrm{Tr}[\hat{\rho}_{\mathcal{O}}\hat{E}_{p_{\mathcal{O}}}(p')]\,\mathrm{d}p'.
\end{aligned}
\tag{50}
$$

Thus the $\mathfrak{p}_{tot}$ measurement can be written as a measurement of the POVM

$$\mathfrak{m} := \{\hat{M}(p)\,\mathrm{d}p\} \qquad \hat{M}(p) = \int_{\mathbb{R}} \lambda(p-p')\hat{E}_{p_{\mathcal{O}}}(p')\,\mathrm{d}p' \tag{51}$$

on the object Hilbert space. If we compare this to (8), we see that

$$\mathfrak{p}_{\mathcal{O}} \to \mathfrak{m}. \tag{52}$$

Hence the measurement of $\mathfrak{p}_{tot}$ can be interpreted as a non-ideal measurement of the object's momentum. Moreover, the inaccuracy of the $\mathfrak{p}_{tot}$-meter as a $\mathfrak{p}_{\mathcal{O}}$-meter is characterized by the initial spread of the $\hat{P}_{\mathcal{A}}$ distribution, $\mathrm{Tr}[\hat{\rho}_{\mathcal{A}}\hat{E}_{p_{\mathcal{A}}}(p)] = \lambda(p)$. Therefore applying (12) leads to:

$$J_{\mathfrak{p}_{\mathcal{O}}\to\mathfrak{m}} = H_{\hat{P}_{\mathcal{A}}} := -\int_{\mathbb{R}} \mathrm{Tr}[\hat{\rho}_{\mathcal{A}}\hat{E}_{p_{\mathcal{A}}}(p)]\log\{\mathrm{Tr}[\hat{\rho}_{\mathcal{A}}\hat{E}_{p_{\mathcal{A}}}(p)]\}\,\mathrm{d}p. \tag{53}$$

Since the apparatus $\mathcal{A}$ measures the POVM $\mathfrak{a}$, this composite scheme would measure $\mathfrak{m}$ and $\mathfrak{a}$ jointly. Furthermore, as $\mathfrak{q}_{\mathcal{O}} \to \mathfrak{a}$, the scheme would in fact be a joint non-ideal measurement of $\mathfrak{q}_{\mathcal{O}}$ and $\mathfrak{p}_{\mathcal{O}}$. But even in such a hypothetical scheme the position and momentum inaccuracies $J_{\mathfrak{q}_{\mathcal{O}}\to\mathfrak{a}}$ and $J_{\mathfrak{p}_{\mathcal{O}}\to\mathfrak{m}}$ cannot violate the uncertainty relation (13). Hence, using (53), it follows that for finite $H_{\hat{P}_{\mathcal{A}}}$ only a less than ideal position measurement is possible [10, 27]:

$$J_{\mathfrak{q}_{\mathcal{O}}\to\mathfrak{a}} \geqslant 1 + \log(\pi) - H_{\hat{P}_{\mathcal{A}}}. \tag{54}$$

In the reasoning that led to (54), we can distinguish three steps. Firstly, it is possible, in principle, to measure a conserved quantity on the meter + object system, after the meter has done its work. Because of the conservation law, this is equivalent to measuring that conserved quantity on the initial meter + object state. Secondly, such a measurement on a composite system may be interpreted as a non-ideal measurement of the corresponding quantity on the object system, with quality determined by the spread of the involved apparatus quantity. Thirdly, the inaccuracy principle comes in, linking the quality of the meter to that of the possible overall measurement, i.e. to the aforementioned spread in the apparatus quantity. We can then try to apply such reasoning to more general cases. As regards the second step, it can be seen [6, 13] that a measurement of an operator $\hat{F}_{tot} = f(\hat{F}_{\mathcal{A}}, \hat{F}_{\mathcal{O}})$, for some function f, can be interpreted as a non-ideal measurement of $\hat{F}_{\mathcal{O}}$, the measurement's quality being related to the $\hat{F}_{\mathcal{A}}$ spread. Thus, suppose that a meter measures some operator $\hat{G}_{\mathcal{O}}$, and let there be an $\hat{F}_{\mathcal{O}}$ such that $[\hat{G}_{\mathcal{O}}, \hat{F}_{\mathcal{O}}]_- \neq \hat{0}$ and $\hat{F}_{tot} = f(\hat{F}_{\mathcal{A}}, \hat{F}_{\mathcal{O}})$ is conserved. Then a result analogous to (54) can be derived from a joint measurement uncertainty relation for $\hat{G}_{\mathcal{O}}$ and $\hat{F}_{\mathcal{O}}$, if available. For finite dimensional PVMs, equation (11) can be used.

We can conclude that, if a conserved quantity exists that does not commute with the measured operator, then the measurement must be a non-ideal one in the sense of section 2, its quality being limited by a suitable generalization of (54). Hence it is also impossible that the measurement is exactly a von Neumann measurement in the sense of (21). This conclusion was reached already by Wigner and by Araki and Yanase [8, 9] on the basis of model studies of angular momentum measurements. Yanase [9] concludes that it is a limitation of quantum measurements in addition to the UP. But, as the above reasoning shows, the joint measurement UP *entails* that limits of type (54) hold in the presence of conservation laws. Accordingly, such limits need not be seen as *extra* limitations on quantum measurements, in addition to the UP. Yanase [9] further shows the possibility of realizing at least an 'approximate' measurement. We shall now also present an example of an approximate measurement, in the sense of non-ideality (6).

Consider an apparatus $\mathscr{A}$ that consists of two subsystems, an ancilla $\mathscr{B}$ and a meter $\mathscr{C}$. $\mathscr{C}$ reads out the POVM $\mathfrak{c}$, a non-ideal measurement of $\mathfrak{q}_{\mathcal{O}} - \mathfrak{q}_{\mathscr{B}}$ on the $\mathcal{O}+\mathscr{B}$ Hilbert space. But the operator $Q_{\mathcal{O}} - Q_{\mathscr{B}}$ commutes with total momentum $\hat{P}_{\text{tot}} = \hat{P}_{\mathcal{O}} + \hat{P}_{\mathscr{B}} + \hat{P}_{\mathscr{C}}$, so that its measurement gives no problems in principle. Therefore $\mathfrak{c}$ may approximate $\mathfrak{q}_{\mathcal{O}} - \mathfrak{q}_{\mathscr{B}}$ arbitrarily well, and we choose $\mathfrak{c} = \mathfrak{q}_{\mathcal{O}} - \mathfrak{q}_{\mathscr{B}}$. A suitable interaction Hamiltonian would be [27, 28]

$$\hat{H}_I = \kappa(\hat{Q}_{\mathcal{O}} - \hat{Q}_{\mathscr{B}})\hat{Z}_{\mathscr{C}} \tag{55}$$

where κ is a coupling constant, and $\hat{Z}_{\mathscr{C}}$ is a $\mathscr{C}$-operator commuting with $\hat{P}_{\mathscr{C}}$. The $(\mathfrak{q}_{\mathcal{O}} - \mathfrak{q}_{\mathscr{B}})$ measurement on the $\mathcal{O}+\mathscr{B}$ Hilbert space can be seen as a measurement of a POVM $\mathfrak{a}$ on the object Hilbert space, $\mathfrak{a}$ being calculated by partial tracing [cf (16)]. The latter POVM is, analogous to (50), a convolution-type non-ideal measurement of $\mathfrak{q}_{\mathcal{O}}$, with quality determined by the initial spread in $\mathfrak{q}_{\mathscr{B}}$,

$$J_{\mathfrak{q}_{\mathcal{O}} \to \mathfrak{a}} = H_{\hat{Q}_{\mathscr{B}}}. \tag{56}$$

We also have an entropic uncertainty relation [19],

$$H_{\hat{Q}_{\mathscr{B}}} + H_{\hat{P}_{\mathscr{B}} + \hat{P}_{\mathscr{C}}} \geqslant 1 + \log(\pi) \tag{57}$$

where we have used the fact that the pair $\{\hat{Q}_{\mathscr{B}}, \hat{P}_{\mathscr{B}} + \hat{P}_{\mathscr{C}}\}$ is unitarily equivalent to $\{\hat{Q}_{\mathscr{B}}, \hat{P}_{\mathscr{B}}\}$. Since $\hat{P}_{\mathscr{A}} = \hat{P}_{\mathscr{B}} + \hat{P}_{\mathscr{C}}$, combining (56) and (57) gives

$$J_{\mathfrak{q}_{\mathcal{O}} \to \mathfrak{a}} \geqslant 1 + \log(\pi) - H_{\hat{P}_{\mathscr{A}}}. \tag{58}$$

Thus, equation (54) is satisfied. A similar model is given by Ozawa [27].

8. Conclusions

One of the interpretations Heisenberg gave the uncertainty principle in his 1927 paper, involves measurement accuracy on the one hand and the disturbance of an incompatible observable on the other. This form of the uncertainty principle can be derived from an uncertainty principle that limits the accuracy achievable in a joint measurement of two incompatible observables. The latter uncertainty principle can also be used to derive bounds for the measurement of observables that do not commute with conserved quantities.

We gave an example which showed that imprecise measurements of such observables are nevertheless possible. Another example, analogous to the γ-microscope, illustrated the disturbance uncertainty principle. Two important properties of the device followed from very natural demands: that it *measures* position follows from the fact that it does not *disturb* position, and that momentum information remains in the momentum distribution, follows from the meter's covariance. Both demands were combined in a concrete scheme, where the position inaccuracy and momentum disturbance follow from two complementary properties of the set of pointer states.

Acknowledgments

The authors thank J Uffink for helpful discussions. They acknowledge the support of the Foundation for Philosophical Research (SWON), which is subsidized by the Netherlands Organization for Scientific Research (NWO).

References

[1] Heisenberg W 1927 *Z. Phys.* **43** 172
[2] Bohr N 1928 *Nature* (suppl.) **121** 580
[3] Ballentine L 1970 *Rev. Mod. Phys.* **42** 358
Popper K 1982 *Quantum Theory and the Schism in Physics* ed W Bartley (Totowa, NJ: Rowman & Littlefield)
[4] Ozawa M 1989 *Squeezed and Non-Classical Light (NATO ASI B 190)* ed P Tombesi and R Pike (New York: Plenum) p 263
[5] Kraus K 1987 *Phys. Rev.* D **35** 3070
[6] Martens H and de Muynck W 1990 *Found. Phys.* **20** 357
[7] Martens H 1991 The uncertainty principle *PhD Thesis* Eindhoven University of Technology
[8] Wigner E 1952 *Z. Phys.* **133** 101
Araki H and Yanase M 1961 *Phys. Rev.* **120** 622
[9] Yanase M 1961 *Phys. Rev.* **123** 666
[10] Busch P 1985 *J. Phys. A: Math. Gen.* **18** 3351
[11] Davies E 1976 *Quantum Theory of Open Systems* (London: Academic)
[12] Holevo A 1982 *Probabilistic and Statistical Aspects of Quantum Mechanics* (Amsterdam: North Holland)
[13] Martens H and de Muynck W 1990 *Found. Phys.* **20** 255
[14] Ali S T and Prugovečki E 1977 *J. Math. Phys.* **18** 219
[15] Schroeck F 1982 *Found. Phys.* **12** 479
[16] Busch P and Lahti P 1984 *Phys. Rev.* D **29** 1634
[17] Uffink J 1991 Private communication
[18] McEliece R 1977 *Theory of Information and Coding* (Reading, MA: Addison-Wesley) ch I
[19] Beckner W 1975 *Ann. Math.* **102** 159
[20] Kraus K 1983 *States, Effects and Operations (Lecture Notes in Physics 190)* (Berlin: Springer)
[21] Lüders G 1951 *Ann. Phys., Lpz.* **8** 322
[22] Martens H and de Muynck W 1991 *Quantum Aspects of Optical Communication (Lecture Notes in Physics 378)* ed C Bendjaballah, O Hirota and S Reynaud (Berlin: Springer) p 41
[23] Hilgevoord J and Uffink J 1991 *Sixty-Two Years of Uncertainty* ed A Miller (New York: Plenum) p 121
[24] Wootters W 1979 *Phys. Rev.* D **19** 473
[25] Uffink J and Hilgevoord J 1988 *Physica* B **51** 309
Hilgevoord J and Uffink J 1991 *Found. Phys.* **21** 323
[26] Martens H 1989 *Phys. Lett.* **137A** 155
[27] Ozawa M 1990 Does a conservation law limit position measurement? *Preprint* Harvard University HUTP-90/B002
[28] Bohm D 1951 *Quantum Theory* (Englewood Cliffs, NJ: Prentice-Hall)