Second International School of Theoretical Physics

SYMMETRY & STRUCTURAL PROPERTIES OF CONDENSED MATTER, *Poznań 1992* (World Sci., Singapore, 1993)

W. Florek, D. Lipiński, T. Lulek, eds. World Scientific Singapore

Quantum Mechanical Observables and Positive Operator Valued Measures

W.M. DE MUYNCK H. MARTENS

Department of Theoretical Physics
Eindhoven University of Technology, Eindhoven, The Netherlands

ABSTRACT

The necessity is discussed of extending the Dirac–von Neumann axiomatization of quantum mechanics to a representation of observables by positive operator valued measures. A partial ordering structure is defined, and a characterization is given of its maximal elements. It is demonstrated that the extension of the formalism allows the joint measurement of incompatible observables, and fully accounts for a mutual disturbance of the measurement results in such measurements. An inequality is presented, illustrating this aspect of quantum mechanical complementarity. The qualitative difference of this inequality with the Heisenberg inequality is discussed. The notion of a Wigner measure is introduced.

1. Introduction

Quantum mechanics is best known in its canonical formulation due to Dirac [1] and von Neumann [2], which can be summarized as follows:

- A state preparation is described by a state function ψ or a density operator ρ.
- A measurement is represented by a self-adjoint operator $A = \int_\Lambda \lambda E(d\lambda)$.
- The expectation value of the measurement outcomes is $\langle A \rangle = \mathrm{Tr}\rho A$.
- The probability distribution of the measurement outcomes is $p(\lambda)d\lambda = \mathrm{Tr}\rho E(d\lambda)$.
- An observable corresponds with a projection valued measure (PVM), defined by the spectral resolution $\{E(d\lambda)\}$ of the self-adjoint operator A, and satisfying:

$$E(d\lambda) \geq 0, \quad E(d\lambda)E(d\lambda') = E(d\lambda)\delta(\lambda - \lambda'), \quad \int_\Lambda E(d\lambda) = I.$$

In recent years, however, it has been realized [3–5] that the Dirac–von Neumann axiomatization offers a too restricted scheme to encompass all measurements that are possible within the domain of quantum mechanics, and that a generalization of this scheme is necessary. The character of this generalization can be derived from a consideration of the measurement process as a quantum mechanical process, in which the interaction between object O and measuring instrument A is governed by a Schrödinger equation in the following way. Let ρ_O and ρ_A be the initial density operators of object and measuring instrument, respectively, and let U represent the unitary operator describing the interaction, then the final state ρ_f is given by

$$\rho_f = U\rho_O \otimes \rho_A U^\dagger.$$

If μ is the variable characterizing the final pointer positions of the measuring instrument, and $\{E^{(A)}(d\mu)\}$ is the PVM corresponding with these pointer states, then the pointer state probabilities are given by

$$p(d\mu) = \mathrm{Tr}\rho_f E^{(A)}(d\mu) = \mathrm{Tr}_O\rho_O \mathrm{Tr}_A\rho_A U^\dagger E^{(A)}(d\mu)U.$$

The set of operators $\{M(d\mu)\}$, defined by

$$p(d\mu) = \mathrm{Tr}_O\rho_O \mathrm{Tr}_A\rho_A U^\dagger E^{(A)}(d\mu)U = \mathrm{Tr}_O\rho_O M(d\mu),$$

consists of the operators describing the relative frequencies of the measurement results of the experiment. It must be compared with the spectral representation $\{E(d\lambda)\}$, which in the Dirac–von Neumann axiomatization plays this role. In general the operators $M(d\mu)$ are not projection operators. They only satisfy

$$M(d\mu) \geq 0, \quad \int_M M(d\mu) = I.$$

It was proven by Holevo [5] that any probability distribution on the set M of pointer positions μ can be represented by the expectation values of the operators $M(d\mu)$. The operators $M(d\mu)$ are said to define *a positive operator valued measure* (POVM). Quantum mechanical observables satisfying the Dirac–von Neumann axiomatization, being represented by PVMs, evidently constitute a subclass of the general class of observables represented by POVMs.

In this contribution we shall first give two examples of measurements needing the generalized formalism for their description. A condition is given for the observable to be related to an observable of the Dirac–von Neumann type. This relation in a natural way gives rise to a definition of a nonideal measurement of an observable, inducing a partial ordering relation on the set of generalized observables. This ordering relation is compared with other ordering definitions that can be found in literature. A quantitative measure is given, expressing a measurement's deviation from ideality. The maximal elements of the ordering turn out to encompass a larger class of observables than the maximal Dirac–von Neumann ones. A characterization of these maximal elements leads to a unified view of orthogonal and nonorthogonal resolutions of the identity.

As an interesting application of the theory of generalized observables the joint nonideal measurement is discussed of observables that are incompatible in the Dirac–von Neumann sense. Two examples are discussed, demonstrating the possibility of such measurements, and showing the failure of the Dirac–von Neumann formalism to describe such measurements. Using the measure of nonideality introduced before, it is possible to give a quantitative characterization of the mutual disturbance of measurement results in such measurements as predicted by Bohr's complementarity principle.

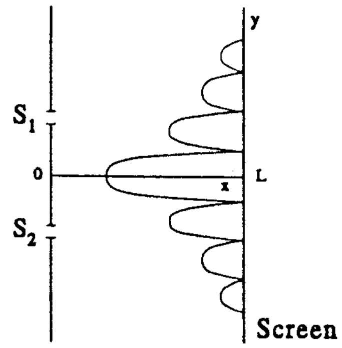

Figure 1: The two-slit experiment

For finite-dimensional Hilbert spaces a general inequality illustrating this complementarity was derived from the general formalism [6]. This inequality is compared with Heisenberg's inequality. Ballentine's criticism [7] as to the latter inequality's failure to describe joint measurement of incompatible observables is endorsed and strengthened. Finally, the joint measurement POVM is shown to be related to an operator valued (Wigner) measure, the expectation values of which defining a quasi-probability distribution analogous to the Wigner distribution.

2. Examples of Generalized Measurements

2.1. The Two-Slit Experiment

The two-slit experiment can be formulated as a problem in a two-dimensional Hilbert space as follows. Let ψ_i, $i = 1, 2$, be two solutions of the Schrödinger equation, corresponding with a particle passing with certainty through slit i. Then, the elements of the two-dimensional Hilbert space are given by

$$\psi = \alpha\psi_1 + \beta\psi_2.$$

The observable measured by this experiment is the position y at which the particle impinges on the screen (cf. Fig. 1). The probability distribution $p(y)$ is represented by a sesquilinear functional

$$p(y) = |\alpha|^2 p_{11}(y) + \alpha^*\beta p_{12}(y) + \alpha\beta^* p_{21}(y) + |\beta|^2 p_{22}(y)$$

on the two-dimensional Hilbert space. It defines a POVM $\{M(dy), -\infty < y < \infty\}$ according to

$$p(y) = \langle \psi | M(y) | \psi \rangle .$$

An important point to be noticed is that the observable measured in this experiment has a *continuous* spectrum $-\infty < y < \infty$. Since no self-adjoint operators on a finite-dimensional Hilbert space exist having a continuous spectrum, it follows that this

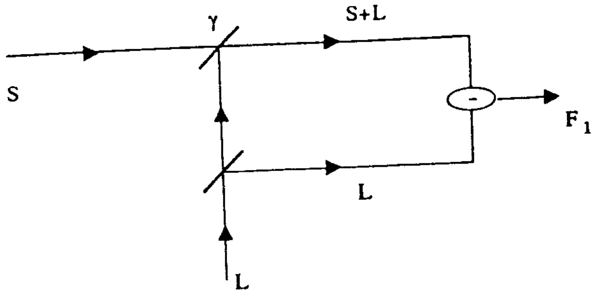

Figure 2: Optical homodyning as a measurement of Q

experiment cannot be described by the conventional Dirac–von Neumann formalism. An extension of the concept of a quantum mechanical observable to a description by positive operator valued measures is seen to be essential in order to be able to encompass also this experiment.

2.2. *Homodyne Optical Detection*

As a second example consider homodyne optical detection (cf. Fig. 2), in which the signal-to-noise ratio of an optical signal S is improved by superposing it with a constant coherent optical signal L produced by a so-called local oscillator. Since $|L| \gg |S|$, the optical intensity F_1 registered by the detector,

$$F_1 = (S+L)^2 - L^2 = S^2 + 2SL \approx 2SL,$$

is seen to be a largely amplified measure of the signal amplitude. By Yuen and Shapiro [8] a quantum mechanical analysis of this experiment has been given for incoming monochromatic signal state $|\psi_S\rangle$. They found the probability distribution

$$p(F_1)dF_1 = \langle\psi_S|\, M_1(dF_1)\, |\psi_S\rangle$$

of the detected intensity F_1 according to

$$M_1(dF_1) = dF_1 \int_{-\infty}^{\infty} dq\ \lambda(F_1 - q)\, |q\rangle\langle q|, \tag{1}$$

$$\lambda(F_1 - q) = \frac{1}{\sqrt{\pi(\gamma^{-1}-1)}} e^{-\frac{(F_1-q)^2}{\gamma^{-1}-1}}. \tag{2}$$

Here $\{|q\rangle\langle q|\, dq\}$ is the spectral representation of the self-adjoint operator $Q = \frac{1}{\sqrt{2}}(a + a^\dagger)$, a being the annihilation operator. It is clear from these equations that $\{M_1(dF_1)\}$ is not a projection valued measure, although in this experiment it is related to one in accordance with (1).

It is interesting to notice that by changing the phase relation between signal and local oscillator beams, as indicated in Fig. 3, by $\pi/2$, the probability distribution of

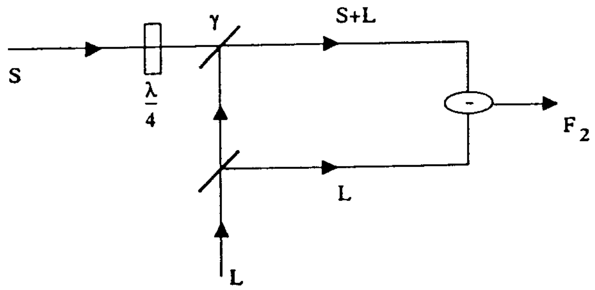

Figure 3: Optical homodyning as a measurement of P

the detected intensity, F_2, takes the form

$$\begin{aligned} p(F_2)dF_2 &= \langle \psi_S | M_2(dF_2) | \psi_S \rangle , \\ M_2(dF_2) &= dF_2 \int_{-\infty}^{\infty} dp \; \mu(F_2 - p) \, |p\rangle \langle p| , \\ \mu(F_2 - p) &= \frac{1}{\sqrt{\pi(\gamma^{-1} - 1)}} \exp\left(-\frac{(F_2 - p)^2}{\gamma^{-1} - 1} \right) . \end{aligned}$$

Here $\{|p\rangle \langle p| \, dp\}$ is the spectral representation of the operator $P = \frac{-i}{\sqrt{2}}(a - a^\dagger)$, complementary with Q. Once again the measurement results do not correspond directly with a PVM but with a POVM.

3. Nonideal Measurements

3.1. Nonideal Measurement of Dirac-von Neumann Observables

If measurement outcomes are described by a POVM that is not a PVM, there need not be any relation with a Dirac-von Neumann observable at all. In certain special cases, like the optical homodyning experiments discussed in Sec. 2.2, however, such a relation turns out to exist (cf. Eq. (1)). Taking expectation values in the left and right hand sides of this equality shows that the probability distribution $\mathrm{Tr}\rho M_1(dF_1)$ is a convolution of the probability distribution $\langle q|\rho|q\rangle$ of the position operator Q with a Gaussian function. Evidently, the homodyning experiment does not yield *ideal* information on the position probability, although, for instance, its mean value is reproduced correctly for any ρ. For this reason the homodyning experiment of Fig. 2 can be interpreted as a *nonideal* measurement of the Dirac-von Neumann position observable. Analogously, the measurement arrangement of Fig. 3 is a nonideal measurement of momentum P.

More generally, we shall say that a measurement procedure is a nonideal measurement of the Dirac-von Neumann observable with PVM $\{F(d\lambda)\}$ if its POVM $\{M(d\mu)\}$

is related to $\{F(d\lambda)\}$ according to

$$M(d\mu) = \int_\Lambda g(d\mu,\lambda)F(d\lambda),\ g(d\mu,\lambda) \geq 0,\ \int_M g(d\mu,\lambda) = 1. \tag{3}$$

From this it is easily seen that a necessary condition for the POVM $\{M(d\mu)\}$ to represent a nonideal measurement of the PVM $\{F(d\lambda)\}$ is $[M(d\mu), F(d\lambda)]_- = 0$. With

$$M(d\mu) = \mathrm{Tr}_A \rho_A U^\dagger E^{(A)}(d\mu)U,\ U = \exp(-iHT/\hbar),$$

it follows that this condition is satisfied if the Hamiltonian H describing the measurement interaction commutes with the PVM $\{F(d\lambda)\}$, thus yielding an operational criterion for devising such a measurement procedure.

3.2. *Partial Ordering of Observables*

In the following we will discuss a generalized notion of nonideal measurement in which the nonideality relation (3) is applied also if $\{F(d\lambda)\}$ is *not* a PVM. By this relation, then, a partial ordering is defined on the set of generalized observables, comparing measurement procedures with respect to their capacity of separating object states. Thus, if $\{F(d\lambda)\}$ and $\{M(d\mu)\}$ are two POVMs related according to the nonideality relation (3), and ρ_1 and ρ_2 represent two object states, then

$$\{\mathrm{Tr}\rho_1 F(d\lambda) = \mathrm{Tr}\rho_2 F(d\lambda)\} \to \{\mathrm{Tr}\rho_1 M(d\mu) = \mathrm{Tr}\rho_2 M(d\mu)\}.$$

Since the converse need not be true, in general $\{F(d\lambda)\}$ has the larger state separation capacity, and, for this reason, can be considered to describe the better measurement of the two.

Before proceeding with the theory of nonideal measurements in the sense of the definition (3), we first want to mention two alternative approaches. The first one [9] is based on the natural partial ordering $A < B$ defined on the set of so-called effect operators A, $0 \leq A \leq I$. In this case A is said to represent a measurement of B, the latter being the less informative one. A second approach is based on the requirement of unbiasedness, to the effect that the POVM $\{M(d\mu)\}$ is thought to represent a measurement of the Dirac–von Neumann observable A if

$$\left\langle \psi \right| \int \mu M(d\mu) \left| \psi \right\rangle = \langle \psi | A | \psi \rangle \tag{4}$$

for any state function $|\psi\rangle$. It was proven in [10] that, on this definition, the standard deviation

$$\Delta(\{M(d\mu)\}) = \int \mu^2 \langle M(d\mu)\rangle - \left\langle \int \mu M(d\mu) \right\rangle^2$$

of the probability distribution generated by $\{M(d\mu)\}$ satisfies the inequality

$$\Delta(\{M(d\mu)\}) \geq \Delta(\{F(d\lambda)\}).$$

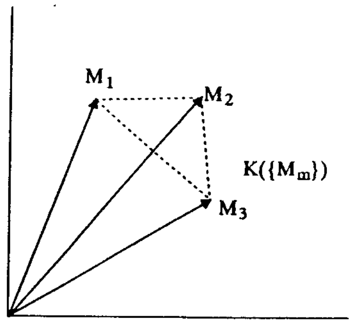

Figure 4: Representation of a POVM

This might be interpreted as a smearing of the ideal probability distribution represented by the PVM $\{F(d\lambda)\}$ due to a disturbing influence of the measurement procedure. Note, however, that our definition (3) of nonideality does not require unbiasedness. Neither is (4) always satisfied in realistic experiments even if a Dirac–von Neumann observable can be defined.

Returning to the ordering based on our nonideality definition, we define equivalence of POVMs as follows. Two POVMs $\{M(d\lambda)\}$ and $\{N(d\mu)\}$ are *equivalent* if they are each other's nonideal measurements, i.e.,

$$M(d\lambda) = \int_M g(d\lambda,\mu)N(d\mu),\ g(d\lambda,\mu) \geq 0,\ \int_\Lambda g(d\lambda,\mu) = 1,$$

and

$$N(d\mu) = \int_\Lambda h(d\mu,\lambda)M(d\lambda),\ h(d\mu,\lambda) \geq 0,\ \int_M h(d\mu,\lambda) = 1.$$

As an example, consider the POVMs $\{M_1, M_2\}$ and $\{N_1, N_2, N_3\}$, both having a discrete spectrum, and satisfying

$$\begin{pmatrix} M_1 \\ M_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} N_1 \\ N_2 \\ N_3 \end{pmatrix},$$

$$\begin{pmatrix} N_1 \\ N_2 \\ N_3 \end{pmatrix} = \begin{pmatrix} \alpha & 0 \\ 1-\alpha & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}, 0 \leq \alpha \leq 1.$$

Equivalent POVMs have equal state separation capacities.

A POVM $\{M(d\lambda)\}$ can be represented by a cone $K(\{M(d\lambda)\})$ in the Banach space of bounded operators [4, 11] (cf. Fig. 4). If POVM $\{M(d\lambda)\}$ represents a nonideal measurement of POVM $\{N(d\mu)\}$ this means that $K(\{N(d\mu)\}) \supseteq K(\{M(d\lambda)\})$ (cf. Fig. 5). In case of equivalence we have $K(\{N(d\mu)\}) = K(\{M(d\lambda)\})$. Intuitively, the information conveyed by a measurement is the more optimal the wider the cone spanned by the operators of its POVM.

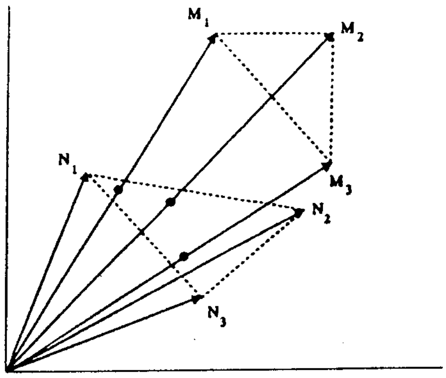

Figure 5: Representation of a nonideality relation

3.3. Maximal POVMs

A POVM $\{M(d\mu)\}$ is maximal if the existence of a POVM $\{N(d\lambda)\}$ satisfying

$$M(d\mu) = \int_\Lambda g(d\mu,\lambda)N(d\lambda),\ g(d\mu,\lambda) \geq 0,\ \int_M g(d\mu,\lambda) = 1,$$

implies equivalence of $\{M(d\mu)\}$ and $\{N(d\lambda)\}$. Hence, the cone of a maximal POVM cannot be included in a still wider cone. It is not difficult to see that the PVM of a maximal Dirac–von Neumann observable is maximal in the sense defined above. Thus, the PVM $\{M_m\} = \{E_m\}, E_m = |m\rangle\langle m|$ a 1-dimensional projection operator on harmonic oscillator eigenstate $|m\rangle$, is a maximal POVM. However, the set of maximal POVMs is much wider than the set of maximal PVMs. It was possible [12] to prove rigorously that on finite-dimensional Hilbert spaces a necessary and sufficient condition for a POVM $\{M_m\}$, having a discrete spectrum, to be maximal is that

$$M_m = c_m E_m,\quad 0 \leq c_m \leq 1,\quad E_m \text{ a 1-dimensional projection operator.}$$

It is not difficult to find examples of such maximal POVMs in which the projection operators E_m do not commute for different values of m. Although not rigorously proven it may safely be conjectured that the POVM $\{\frac{d^2\alpha}{\pi}|\alpha\rangle\langle\alpha|\}$, $|\alpha\rangle$ a coherent state, is maximal. It is interesting to note that, from the point of view of maximality as defined above, there is no difference between orthogonal and nonorthogonal resolutions of the identity.

By Ludwig [4] an alternative notion of nonideality is considered for measurement procedures, leading to a different sense in which these procedures are ordered. In Ludwig's approach nonideal measurements are represented by convex combinations $\{\lambda M_k + (1-\lambda)N_k\}$, $0 \leq \lambda \leq 1$ of the POVMs $\{M_k\}$ and $\{N_k\}$. The optimal elements of this ordering are now taken as the extremal elements of the convex set thus obtained. Drawbacks of this approach are:

- On the one hand the possibility of convex combination of POVMs is too restricted because only POVMs having the same spectra can be combined in this way.
- On the other hand the possibility of convex combination is too liberal because *any* pair of POVMs having equal spectra can be combined, even pairs lacking any physical relation as to their physical meaning (like, e.g., position and momentum).
- There is no relation to information content: a PVM $\{E_k\}$ may contain multidimensional projectors and nevertheless be extremal in the sense of the convexity approach.

4. Information and Generalized Quantum Mechanics

4.1. Informational (In)Completeness of Maximal POVMs

As is well known in classical mechanics there is one unique maximal measure on phase space, viz.,

$$\{M_{\omega'}(d\omega)\} = \{\delta(\omega' - \omega)d\omega\}, \quad \omega', \omega \in \text{phase space},$$

yielding complete information on the (statistical) state of the object. Any measurement yielding more coarse-grained information is a nonideal measurement of this maximal observable.

In the quantum mechanics of Dirac and von Neumann the situation is quite different. If $\{\alpha_k\}$ and $\{\beta_l\}$ are two incompatible complete orthonormal sets of functions, then $\{E_k\} = \{|\alpha_k\rangle\langle\alpha_k|\}$ and $\{F_l\} = \{|\beta_l\rangle\langle\beta_l|\}$ are both maximal PVMs. Their expectation values, however, only contain partial information on the quantum mechanical state function, the information of the PVMs moreover being inequivalent because of incompatibility. Bohr's complementarity principle has largely contributed to the idea that this is the essential difference between classical and quantum mechanics: on measuring one observable the other (incompatible) one must be disturbed so as to wipe out information on the latter observable. For this reason it should be impossible to obtain *complete* information on the quantum mechanical state function by means of the measurement of a single observable.

It seems that this difference between classical mechanics and quantum mechanics becomes less essential when the possibility is taken into account of measurement procedures described by POVMs. For instance, as is well known the maximal POVM $\{\frac{d^2\alpha}{\pi}|\alpha\rangle\langle\alpha|\}$ yields *complete* information on the state function. Hence, if it would be possible to find a measurement procedure described by this POVM, this procedure would constitute a complete measurement in this sense. It was demonstrated by Busch [13] that such complete measurements are actually possible. In the following we will discuss a quantum optical measurement procedure that is already widely used in actual practice, satisfying this completeness criterion.

It is clear from this that the theory of generalized observables describes a type of measurements essentially different from the ones described by the Dirac–von Neumann formalism.

4.2. A Measure of Nonideality

In this section we will restrict ourselves to finite-dimensional Hilbert spaces and finite spectra. A generalization to infinite discrete spectra seems to be straightforward; continuous spectra may need more sophisticated mathematical methods. For simplicity we shall also restrict to nonideal measurements of maximal PVMs. Let $\{M_{m'}\}_{m'=1}^{N'}$ represent a nonideal measurement of the maximal PVM $\{E_m\}_{m=1}^{N}$. Then

$$M_{m'} = \sum_{m=1}^{N} g_{m'm} E_m, \; g_{m'm} \geq 0, \; \sum_{m'=1}^{N'} g_{m'm} = 1. \tag{5}$$

As a measure of nonideality of the nonideality matrix $(g_{m'm})$ a number of quantities are useful, like e.g., Shannon's channel capacity [14], being a measure for the quality of a transmission channel analogous to the one obtained in a nonideal measurement. For our purpose the average row entropy

$$J_{(g)} = -\frac{1}{N} \sum_{m',m} g_{m'm} \log \frac{g_{m'm}}{\sum_l g_{m'l}} \tag{6}$$

of the matrix $(g_{m'm})$ is more suitable. It has the following properties [12, 14]:

- $0 \leq J_{(g)} \leq \log N$, if $N \leq N'$, $\log \frac{N}{N'} \leq J_{(g)} \leq \log N$, if $N > N'$;
- $J_{(g)} = 0$ if $g_{m'm} = \delta_{m'm}$;
- $J_{(g)} = \log N$ if $g_{m'm} = \frac{1}{N'}$;
- If (g) and (h) are stochastic matrices, satisfying the conditions mentioned in (5), then $J_{(hg)} \geq J_{(g)}$.

Hence, this quantity becomes larger as the measurement becomes more nonideal in the sense of blurring information by allowing more interaction between the channels.

5. Joint Measurement of Incompatible Observables

5.1. Definitions

We shall first define *commensurability* of two observables, i.e., specify conditions under which two observables can be measured jointly. Essential is that a measurement setup, achieving this, must have *two* pointers, one for each observable. The joint probability distribution of the two pointers must be specified by a bivariate POVM. Hence, we arrive at the following definition:
The observables represented by POVMs $\{M(d\mu)\}$ and $\{N(d\nu)\}$ are commensurable if a measurement setup exists, the probabilities of which are determined by the bivariate POVM $\{R(d\mu, d\nu)\}$, satisfying

$$M(d\mu) = R(d\mu, \int d\nu),$$

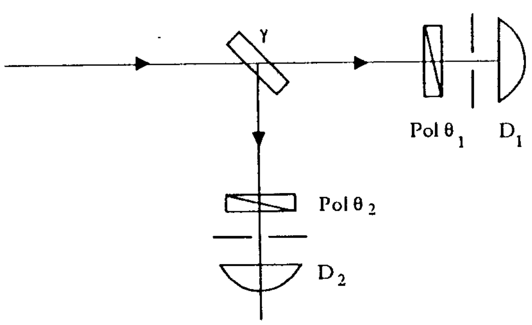

Figure 6: Joint nonideal measurement of two polarization observables

$$N(d\nu) = R(\int d\mu, d\nu).$$

The following theorem was proven [15, 16]:
If $\{M(d\mu)\}$ and $\{N(d\nu)\}$ are commensurable PVMs, then

$$[M(d\mu), N(d\nu)]_- = 0.$$

As follows from this theorem, observables represented by incompatible PVMs cannot be commensurable. This corresponds with the impossibility of measuring jointly ('simultaneously') incompatible observables as a well-known consequence of the Dirac–von Neumann formalism of quantum mechanics.

It turns out that the generalization of the formalism to encompass observables being described by POVMs can make use of the fact that such generalized observables can be commensurable without commutativity of the POVMs. More specifically, the commensurable POVMs may describe *nonideal* measurements of incompatible Dirac–von Neumann observables. This can be summarized in the following definition of *a joint nonideal measurement*:
The observables described by POVMs $\{M(d\mu)\}$ and $\{N(d\nu)\}$ are jointly nonideally measurable if a measurement setup exists having a bivariate POVM $\{R(d\mu, d\nu)\}$, satisfying:

$$\begin{aligned} R(d\mu, \int d\nu) &= \int_{\mu'} g(d\mu, \mu')M(d\mu'), \; g(d\mu, \mu') \geq 0, \; g(\int d\mu, \mu') = 1, \\ R(\int d\mu, d\nu) &= \int_{\nu'} h(d\nu, \nu')N(d\nu'), \; h(d\nu, \nu') \geq 0, \; h(\int d\nu, \nu') = 1. \end{aligned} \quad (7)$$

In the following subsections we shall describe examples of nonideal measurements in which $\{M(d\mu)\}$ and $\{N(d\nu)\}$ are PVMs.

5.2. Joint Measurement of Incompatible Polarization Observables

By the measurement setup of Fig. 6 two incompatible polarization observables of a photon are measured jointly in the sense described in Sec. 5.1. The setup consists of a beam splitter with transmission probability γ, and polarizers in directions θ_1 and

θ_2 in each of the two outgoing beams, respectively. We shall denote the observables of polarization in directions θ_1 and θ_2 as $\{E_+, E_-\}$ and $\{F_+, F_-\}$, respectively. Detectors D_1 and D_2 constitute the two pointers necessary for the setup to perform a joint measurement of two observables. Denoting the event of detecting an incoming photon by $+$, and a failure by $-$, the joint detection probability distribution p_{mn}, $m, n = +$ or $-$, can easily be found according to

$$\begin{aligned} p_{++} &= 0, \\ p_{+-} &= \gamma \langle E_+ \rangle, \\ p_{-+} &= (1-\gamma)\langle F_+ \rangle, \\ p_{--} &= 1 - \gamma\langle E_+ \rangle - (1-\gamma)\langle F_+ \rangle . \end{aligned}$$

Here the expectation values are taken in the initial (incoming) state of the photon. The POVM $\{R_{mn}\}$ describing the experiment now follows directly from

$$p_{mn} = \langle R_{mn} \rangle ,$$

yielding

$$(R_{mn}) = \begin{pmatrix} 0 & \gamma E_+ \\ (1-\gamma)F_+ & 1 - \gamma E_+ - (1-\gamma)F_+ \end{pmatrix} . \tag{8}$$

It is straightforward to calculate the marginals of the matrix (R_{mn}). These can be written down in matrix notation according to

$$\begin{pmatrix} \sum_n R_{+n} \\ \sum_n R_{-n} \end{pmatrix} = \begin{pmatrix} \gamma & 0 \\ 1-\gamma & 1 \end{pmatrix} \begin{pmatrix} E_+ \\ E_- \end{pmatrix},$$

$$\begin{pmatrix} \sum_m R_{m+} \\ \sum_m R_{m-} \end{pmatrix} = \begin{pmatrix} 1-\gamma & 0 \\ \gamma & 1 \end{pmatrix} \begin{pmatrix} F_+ \\ F_- \end{pmatrix}.$$

Evidently this precisely satisfies the definition (7) of a joint nonideal measurement of $\{E_+, E_-\}$ and $\{F_+, F_-\}$.

5.3. *Four-Port Optical Homodyning as a Joint Nonideal Measurement of Q and P*

In the four-port homodyning setup the two homodyning setups discussed in Sec. 2.2 are combined in the way exhibited in Fig. 7, γ,ϵ_1 and ϵ_2 being transmission probabilities of semi-transparent mirrors. The joint probability distribution of the two output intensities F_1 and F_2 is given by [8]

$$p(dF_1, dF_2) = \langle \psi_S | R(dF_1, dF_2) | \psi_S \rangle ,$$

$$R(dF_1, dF_2) = \frac{dF_1 dF_2}{\pi \sqrt{(\frac{1}{(1-\gamma)\epsilon_1} - 2)(\frac{1}{\gamma\epsilon_2} - 2)}} \int \frac{d^2\beta}{\pi} e^{-\frac{(F_1 - \sqrt{2}\,\mathrm{Re}\beta)^2}{\frac{1}{(1-\gamma)\epsilon_1} - 2} - \frac{(F_2 - \sqrt{2}\,\mathrm{Im}\beta)^2}{\frac{1}{\gamma\epsilon_2} - 2}} |\beta\rangle \langle\beta| .$$

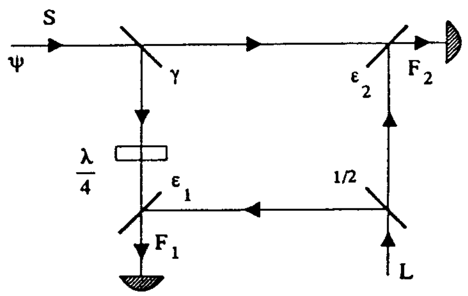

Figure 7: Four-port optical homodyning as a joint nonideal measurement of Q and P

Here $|\beta\rangle$ is a coherent state. Calculating the two marginals of the bivariate POVM $\{R(dF_1, dF_2)\}$ we find

$$R(dF_1, \int_{-\infty}^{\infty} dF_2) = dF_1 \int_{-\infty}^{\infty} dq\; g(F_1 - q)\,|q\rangle\langle q|, \tag{9}$$

$$R(\int_{-\infty}^{\infty} dF_1, dF_2) = dF_2 \int_{-\infty}^{\infty} dp\; h(F_2 - p)\,|p\rangle\langle p|, \tag{10}$$

with:

$$g(F_1 - q) = \frac{1}{\delta_1\sqrt{\pi}} \exp\left(-\frac{(F_1 - q)^2}{\delta_1^2}\right),\ \delta_1^2 = \frac{1}{(1-\gamma)\epsilon_1} - 1, \tag{11}$$

$$h(F_2 - p) = \frac{1}{\delta_2\sqrt{\pi}} \exp\left(-\frac{(F_2 - p)^2}{\delta_2^2}\right),\ \delta_2^2 = \frac{1}{\gamma\epsilon_2} - 1. \tag{12}$$

Once again we have a joint nonideal measurement of two incompatible observables, viz., the observables Q and P defined in Sec. 2.2.

6. Complementarity

6.1. Examples

The two nonideality matrices

$$(g) = \begin{pmatrix} \gamma & 0 \\ 1-\gamma & 1 \end{pmatrix},\quad (h) = \begin{pmatrix} 1-\gamma & 0 \\ \gamma & 1 \end{pmatrix} \tag{13}$$

of Sec. 5.2 for the measurement of polarization observables in directions θ_1 and θ_2, respectively, have a complementary behavior in the sense that their dependence on the parameter γ is opposite. Thus, in the limit $\gamma = 1$ we have

$$(g) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\quad (h) = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \tag{14}$$

implying the polarization measurement in direction θ_1 to be an ideal one. However, in this limit the other marginal reduces to the POVM $\{0, I\}$. Hence, as is also to be expected from the fact that in this limit no photons reach the polarizer θ_2, no information is obtained about the polarization observable in this latter direction.

In the limit $\gamma = 0$ we get

$$(g) = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \quad (h) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Hence, in this limit it is just the other way around: the polarization measurement in direction θ_2 is ideal, whereas the other one is uninformative. The two limits evidently correspond with two mutually excluding measurement arrangements of two incompatible Dirac–von Neumann observables, in which information on the other, incompatible, one is wiped out completely. For intermediate values of γ, satisfying $0 < \gamma < 1$, information is obtained on *both* observables, information on one observable being more deteriorated as information on the other one is of higher quality. This can be interpreted as a mutual disturbance in the joint measurement of incompatible observables: evidently the quality of information on one observable is influenced in a negative sense by the presence of the part of the measurement arrangement that is necessary for obtaining information on the other observable.

Using the measure of nonideality introduced in Sec. 4.2 this notion of complementarity can be given a quantitative expression. Applying (6) to the nonideality matrices (13) we obtain

$$J_{(g)} = \frac{1}{2}[(2-\gamma)\log(2-\gamma) - (1-\gamma)\log(1-\gamma)], \qquad (15)$$

$$J_{(h)} = \frac{1}{2}[(1+\gamma)\log(1+\gamma) - \gamma\log\gamma], \qquad 0 \leq \gamma \leq 1. \qquad (16)$$

Eliminating γ from (15) and (16) we obtain the graph depicted in Fig. 8. It is easily seen that the nonideality measures $J_{(g)}$ and $J_{(h)}$ satisfy the inequality

$$J_{(g)} + J_{(h)} \geq \log 2.$$

From this result it is clear that $J_{(g)} = J_{(h)} = 0$ is impossible, once more expressing the fact that the incompatible polarization observables cannot *both* be measured in an ideal way by this measurement arrangement.

With respect to complementarity the situation in the four-port homodyning example of Sec. 5.3 is completely analogous. Taking the limit $\gamma = 0$, $\epsilon_1 = 1$, the nonideality functions (11) and (12) reduce to

$$g(F_1 - q) = \delta(F_1 - q), \quad h(F_2 - p) = 0,$$

making the Q measurement ideal, and the P measurement uninformative. On the other hand, in the limit $\gamma = 1$, $\epsilon_2 = 1$ we get

$$g(F_1 - q) = 0, \quad h(F_2 - p) = \delta(F_2 - p),$$

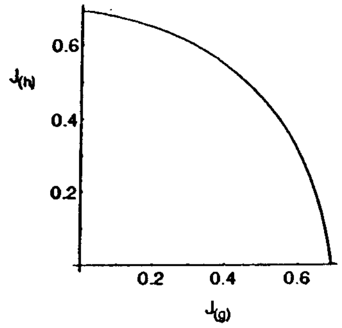

Figure 8: Complementarity in the joint nonideal measurement of two polarization observables

thus realizing the complementary measurement setup in which the Q measurement is uninformative, and the P measurement ideal. Once again for intermediate values of γ the experiment yields information on *both* Q and P, although this information is subject to a deterioration due to the mutual disturbance induced by combining the two measurement procedures of Sec. 2.2 in one single measurement arrangement. Since

$$\delta_1\delta_2 = \sqrt{\left(\frac{1}{(1-\gamma)\epsilon_1} - 1\right)\left(\frac{1}{\gamma\epsilon_2} - 1\right)} \geq \sqrt{\left(\frac{1}{1-\gamma} - 1\right)\left(\frac{1}{\gamma} - 1\right)} = 1,$$

it is clear that $\delta_1 = \delta_2 = 0$ is impossible. Hence, also here the measurement arrangement does not allow an ideal measurement of both incompatible observables.

6.2. A General Inequality

In [6] the following theorem was rigorously proven:
If $\{R_{mn}\}$ describes a joint nonideal measurement of two observables represented by PVMs $\{M_m\}$ and $\{N_n\}$ on a finite-dimensional Hilbert space, i.e.,

$$\sum_n R_{mn} = \sum_{m'} g_{mm'} M_{m'},\ g_{mm'} \geq 0,\ \sum_m g_{mm'} = 1,$$

$$\sum_m R_{mn} = \sum_{n'} h_{nn'} N_{n'},\ h_{nn'} \geq 0,\ \sum_n h_{nn'} = 1,$$

and $J_{(g)}$ is the nonideality measure defined by (6) (and analogously for (h)), then

$$J_{(g)} + J_{(h)} \geq -\log\{\max_{m,n} \mathrm{Tr} M_m N_n\}. \tag{17}$$

The proof of this theorem follows straightforwardly from the inequality

$$H_{\{M_m\}}(\rho) + H_{\{N_n\}}(\rho) \geq -\log(\max_{m,n} \mathrm{Tr} M_m N_n),$$

found by Maassen and Uffink [17] for the entropy functions $H_{\{M_m\}}(\rho)$ and $H_{\{N_n\}}(\rho)$ defined by

$$H_{\{M_m\}}(\rho) = -\sum_{m'} \langle m'|\rho|m'\rangle \log \langle m'|\rho|m'\rangle$$

(and analogously for $\{N_n\}$). Here $\{|m'\rangle\}$ is the complete orthonormal set of state functions corresponding with observable $\{M_m\}$, and ρ is a density operator satisfying $\mathrm{Tr}\rho = 1$. The proof is based on the fact that the nonideality measure is related to the entropy function according to

$$J_{(g)} = \frac{1}{N}\sum_m(\mathrm{Tr}\sum_n R_{mn})H_{\{M_m\}}\left(\frac{\sum_n R_{mn}}{\mathrm{Tr}\sum_n R_{mn}}\right).$$

Then, because of the convexity of the entropy function, the following inequality can easily be derived:

$$J_{(g)} = \frac{1}{N}\sum_m(\mathrm{Tr}\sum_{n'} R_{mn'})H_{\{M_m\}}\left(\sum_n r_{mn}\frac{R_{mn}}{\mathrm{Tr}R_{mn}}\right) \geq \frac{1}{N}\sum_{mn}(\mathrm{Tr}R_{mn})H_{\{M_m\}}\left(\frac{R_{mn}}{\mathrm{Tr}R_{mn}}\right),$$

in which $r_{mn} = \mathrm{Tr}R_{mn}/\mathrm{Tr}\sum_{n'} R_{mn'}$. Since for $J_{(h)}$ an analogous inequality holds, the Maassen–Uffink inequality can now be applied straightforwardly to yield (17). Hence, if the sets $\{|m\rangle\}$ and $\{|n\rangle\}$ have no state function in common, we have $\mathrm{Tr}M_mN_n < 1$ for all m, n and $J_{(g)} + J_{(h)} > 0$.

If $\mathrm{Tr}M_mN_n = 1$ for some M_m and N_n, then the inequality (17) is fulfilled trivially and does not yield any information on the measurement process. In this case it is possible, however, to sharpen the inequality in the following way. Let $\{G_i\}$ be a PVM defining the common eigenspaces of the PVMs $\{M_m\}$ and $\{N_n\}$. Thus,

$$[G_i, M_m]_- = [G_i, N_n]_- = 0.$$

Then it is possible to prove the inequality [6]

$$J_{(g)} + J_{(h)} \geq \sum_i(\mathrm{Tr}G_i)c_i, \tag{18}$$

$$c_i = -\frac{2}{N}\log[\frac{1}{2} + \frac{1}{2}\max_{m,n}\|M_mN_nG_i\|].$$

This demonstrates that the inequality (18) is nontrivial as long as not all operators M_m commute with all N_n.

7. Connection with the Heisenberg Uncertainty Relations

The idea of complementarity as being related to a mutual disturbance in the simultaneous measurement of incompatible observables has been investigated already by Bohr and Heisenberg in their discussions of the two-slit experiment and the γ microscope [18, 19]. The results obtained in the previous sections are completely in accordance with the conclusions reached by Bohr and Heisenberg. The inequalities (17) and (18) yield perfect descriptions of the uncertainty principle as illustrated, e.g., by Heisenberg's γ microscope. The examples discussed in Secs. 5 and 6 are of the same type.

It is rather generally accepted (e.g. Merzbacher [20]) that complementarity as discussed here is taken into account within the formalism of quantum mechanics by the (Heisenberg) uncertainty relations

$$\Delta A \Delta B \geq \frac{1}{2}\left|\langle [A,B]_- \rangle\right|, \tag{19}$$

$$(\Delta A)^2 = \mathrm{Tr}\rho(A - \langle A\rangle)^2,\ (\Delta B)^2 = \mathrm{Tr}\rho(B - \langle B\rangle)^2. \tag{20}$$

Yet, this interpretation was challenged by Ballentine [7], who objected that the inequality (19) does not refer to the joint measurement of observables A and B, but can be tested by measuring A and B *separately*. Indeed, the standard deviations (20) are calculated using the *projection valued measures* generated by the spectral representations of the self-adjoint operators A and B. No disturbance or nonideality seems to be involved in the measurement if the observables are maximal. For this reason our theory seems to endorse Ballentine's idea that the (Heisenberg) uncertainty relations (19) refer to the object rather than to the measurement, and just express the impossibility of *preparing* a quantum mechanical ensemble that is dispersionless in both of two incompatible (Dirac-von Neumann) observables. Hence, the Heisenberg inequalities refer to joint *preparation* rather than to the joint *measurement* of incompatible quantities.

According to Ballentine quantum mechanics has nothing to say about the joint measurement of incompatible observables. As we have seen this judgment is too pessimistic, although understandable because it is based on the restricted Dirac-von Neumann formalism. As we have seen, the generalized formalism of POVMs *can* describe joint nonideal measurements of incompatible observables. The inequalities (17) and (18) are expressions of complementarity and mutual disturbance in the joint measurement of incompatible observables. It is interesting to notice that these inequalities refer in an unambiguous way to the measurement *only*, since they are completely independent of the density operator. Hence, they do not refer in any way to state preparation, but exhibit properties of the measurement. Since the Heisenberg inequalities refer to both the density operator and the observables A and B, we see that inequality (19) is of a qualitatively different nature. It appears that the uncertainty principle is actually a twofold one, one for preparation described by, e.g., the Heisenberg inequality, and one for measurement described by, e.g., the inequality (17). In the past this distinction has seldom been made, an omission that can be attributed to a failure to see the restricted applicability of the Dirac-von Neumann formalism of quantum mechanics.

8. Wigner Measures

The nonideality matrices (13) can be inverted, yielding

$$(g^{-1}) = \begin{pmatrix} \gamma^{-1} & 0 \\ 1-\gamma^{-1} & 1 \end{pmatrix},\ (h^{-1}) = \begin{pmatrix} (1-\gamma)^{-1} & 0 \\ 1-(1-\gamma)^{-1} & 1 \end{pmatrix}. \tag{21}$$

Using these inverse matrices the following operators can be calculated from the POVM (8):

$$W_{kl} = \sum_{mn} g_{km}^{-1} h_{ln}^{-1} R_{mn}. \tag{22}$$

Since

$$\sum_k g_{km}^{-1} = \sum_l h_{ln}^{-1} = 1,$$

the set of operators $\{W_{kl}\}$ has the following properties:

$$\sum_{kl} W_{kl} = I, \tag{23}$$

$$\sum_l W_{kl} = E_k, \tag{24}$$

$$\sum_k W_{kl} = F_l. \tag{25}$$

Evidently the marginals of the matrix (W_{kl}) define the probability distributions of the two observables $\{E_+, E_-\}$ and $\{F_+, F_-\}$, which, consequently, can be *calculated exactly* from the joint probability distribution p_{mn} defined by $\{R_{mn}\}$. Because of (23) the set of operators $\{W_{kl}\}$ constitutes an operator valued measure. This measure will be called a *Wigner measure*. According to a theorem due to Wigner [21] measures satisfying the marginality relations (24) and (25) cannot be POVMs. Indeed, not all operators W_{kl} defined by (22),

$$(W_{kl}) = \begin{pmatrix} 0 & E_+ \\ F_+ & E_- - F_+ \end{pmatrix}, \tag{26}$$

are positive operators.

For the four-port optical homodyning example of Sec. 5.3 it is possible to calculate a Wigner measure in a completely analogous manner. In the first place the convolution relations (9) and (10) can be inverted, yielding

$$g^{-1}(q - F_1) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{\frac{k^2 \delta_1^2}{4}} e^{ik(q-F_1)},$$

$$h^{-1}(p - F_2) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{\frac{k^2 \delta_2^2}{4}} e^{ik(p-F_2)}.$$

From this result it is straightforward to find the Wigner measure according to

$$\begin{aligned} W(q,p) &= \int\int g^{-1}(q - F_1)\, h^{-1}(p - F_2)\, R(dF_1, dF_2) \\ &= \left(\frac{1}{2\pi}\right)^2 \int_{-\infty}^{\infty} dk \int_{-\infty}^{\infty} dk' \int \frac{d^2\beta}{\pi} |\beta\rangle \langle\beta| \, e^{\frac{k^2+k'^2}{4} + ik(q-\sqrt{2}Re\beta) + ik'(p-\sqrt{2}Im\beta)}, \end{aligned}$$

satisfying

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(q,p) dq dp = I,$$

$$\int_{-\infty}^{\infty} W(q,p)dp = |q\rangle \langle q| ,$$

$$\int_{-\infty}^{\infty} W(q,p)dq = |p\rangle \langle p| .$$

The operator valued measure $\{W(q,p)\}$ yields for the state function $|\psi\rangle$ the (quasi)probability distribution

$$\langle\psi|W(q,p)|\psi\rangle = \frac{1}{\pi\hbar}\int_{-\infty}^{\infty} dy\psi(q+y)^{*}\psi(q-y)e^{\frac{2ipy}{\hbar}} .$$

This is the well-known Wigner distribution [22].

The four-port homodyning example is especially interesting because it follows that, at least in principle, from its results the Wigner distribution can be calculated. Since the Wigner distribution and the state function are informationally equivalent, it follows that the state function $|\psi\rangle$ can be determined completely by this measurement. Such measurements are called *complete measurements* [13]. Since no Dirac–von Neumann observable can yield a complete determination of the state function, this shows clearly the fundamental improvement in our measurement potential in generalizing to measurements described by POVMs.

Not all such measurements are complete measurements, however. Thus, the joint measurement of polarization observables discussed in Sec. 5.2 is not complete. From (8) it is seen that each operator R_{mn} is a linear combination of operators from $\{E_+, E_-\}$ and $\{F_+, F_-\}$. Hence, the joint measurement does not yield more information than the joint information obtained by measuring each of these observables separately (for this reason such a joint measurement is called a *trivial* joint measurement). It is, however, not difficult to device also in the case of polarization observables measurement arrangements yielding complete information on the polarization state of a photon.

Acknowledgment

One of the authors (HM) acknowledges financial support from the Foundation for Philosophical Research (SWON), which is subsidized by the Netherlands Organization for Scientific Research (NWO).

References

1. P.A.M. Dirac, *The Principles of Quantum Mechanics* (4th rev.ed., Clarendon, Oxford, 1967).
2. J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (University Press, Princeton, 1955).
3. E.B. Davies, *Quantum Theory of Open Systems* (Academic Press, London, 1976).
4. G. Ludwig, *Foundations of Quantum Mechanics* (Springer, Berlin, 1983) Vol. I.

5. A.S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).

6. H. Martens and W.M. de Muynck, *Found. of Phys.* **20** (1990) 357.

7. L.E. Ballentine, *Rev. Mod. Phys.* **42** (1970) 358.

8. H.P. Yuen and J. Shapiro, *IEEE Trans. Inf. Theor.* **IT-26** (1980) 78.

9. P. Busch and P.J. Lahti, *Phys. Rev.* **D29** (1984) 1634.

10. P. Kruszynski and W.M. de Muynck, *J. Math. Phys.* **28** (1987) 1761.

11. G. Jameson, *Ordered Linear Spaces* (Springer, New York, 1970).

12. H. Martens and W.M. de Muynck, *Found. of Phys.* **20** (1990) 255.

13. P. Busch, *Found. of Phys.* **17** (1987) 905.

14. R. McEliece, *The Theory of Information and Coding* (Addison–Wesley, London, 1977).

15. W.M. de Muynck and J.P.H.W. van den Eijnde, *Found. of Phys.* **14** (1984) 111.

16. W.M. de Muynck and J.M.V.A. Koelman, *Phys. Lett.* **A98** (1983) 1.

17. H. Maassen and J.B.M. Uffink, *Phys. Rev. Lett.* **60** (1988) 1103.

18. N. Bohr, in *Albert Einstein: Philosopher-Scientist*, ed. P.A. Schilpp (The Library of Living Philosophers, 1949) p. 199.

19. W. Heisenberg, *Die Physikalische Prinzipien der Quantentheorie* (Bibliographisches Institut, Mannheim, 1958).

20. E. Merzbacher, *Quantum Mechanics* (Wiley, New York, 1970).

21. E. Wigner, in *Perspectives in Quantum Theory*, eds. W. Yourgrau and A. van der Merwe (MIT Press, Cambridge, 1971) p. 25.

22. E. Wigner, *Phys. Rev.* **40** (1932) 749.