

An alternative to the Lüders generalization of the von Neumann projection, and its interpretation

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

1998 J. Phys. A: Math. Gen. 31 431

(<http://iopscience.iop.org/0305-4470/31/2/005>)

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

Download details:

IP Address: 171.66.16.122

The article was downloaded on 02/06/2010 at 06:51

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

An alternative to the Lüders generalization of the von Neumann projection, and its interpretation

Willem M de Muynck

† Department of Theoretical Physics, Eindhoven University of Technology, Eindhoven, the Netherlands

Received 7 July 1997, in final form 2 October 1997

Abstract. An alternative to the Lüders generalization of the von Neumann projection is formulated allowing a generalization to measurements represented by positive operator-valued measures. This generalization is based on the notions of non-orthogonal projections and bi-orthonormal systems in Hilbert–Schmidt space, and is closely related to phase-space representations of quantum mechanics. A number of applications will be discussed. It is demonstrated that the alternative projection has a natural interpretation as a representation of the information on the initial state, to be obtained by the measurement.

1. Introduction

The von Neumann–Lüders projection has its origin in the prehistory of quantum measurement theory. It yields a prescription of the way a quantum-mechanical object is ideally influenced by a measurement. For an observable A having a nondegenerate spectrum this was thought by von Neumann [1] to take place according to the scheme

$$|\Psi\rangle = |\psi\rangle \otimes |\theta_0\rangle \rightarrow |\Psi_f\rangle = \sum_m c_m |a_m\rangle \otimes |\theta_m\rangle \quad (1)$$

in which $|\psi\rangle = \sum_m c_m |a_m\rangle$ and $|\theta_0\rangle$ are the initial states of the object and measuring instrument, respectively, $|a_m\rangle$ is the eigenvector of observable A corresponding to eigenvalue a_m , and $|\theta_m\rangle$ is a pointer state of the measuring instrument. Measurements satisfying this scheme are called *measurements of the first kind*, such measurements leaving the object state unchanged if its initial state is an eigenvector of the measured observable.

For the density operator of the object this implies the following transition from initial state ρ to final state ρ_{of} :

$$\rho = |\psi\rangle\langle\psi| \rightarrow \rho_{of} = \text{Tr}_a |\Psi_f\rangle\langle\Psi_f| = \sum_m |c_m|^2 |a_m\rangle\langle a_m|. \quad (2)$$

It is important to note that this ρ_{of} can be written as

$$\rho_{of} = \sum_m (\text{Tr } \rho E_m) E_m \quad (3)$$

where $E_m = |a_m\rangle\langle a_m|$ is the one-dimensional projection operator onto the vector $|a_m\rangle$. Since $\sum_m (\text{Tr } \rho_{of} E_m) E_m = \rho_{of}$ the state transformation (2) is a projection.

In order to accommodate possible degeneracy of the eigenvalues a_m this scheme was generalized by Lüders [2] according to

$$\rho \rightarrow \rho_{of} = \sum_m E_m \rho E_m \quad (4)$$

in which E_m is now the (possibly multidimensional) projection operator of the subspace spanned by the eigenvectors corresponding to a_m . In contrast to the transition to (3), transition (4) is also a projection for multidimensional projection operators E_m . It reduces to the von Neumann projection (2) in the case of non-degeneracy. Apart from being projections, transitions (2) and (4) map density operators onto density operators (i.e. $\rho_{of} \geq O$ if $\rho \geq O$, and $\text{Tr } \rho_{of} = \text{Tr } \rho$). By Ludwig [3] such mappings are referred to as *mimorphisms*.

It is often considered as a virtue that the Lüders projection (4) satisfies the minimum-disturbance property characteristic of first-kind measurements (e.g. Furry [4]). It is easily seen that for both the von Neumann and the Lüders projection we have

$$\text{Tr } \rho_{of} E_m = \text{Tr } \rho E_m \quad (5)$$

indicating that a measurement of observable A in the final object state ρ_{of} yields the same probability distribution as the measurement performed in the initial state ρ . The property (5) is sometimes taken as defining a measurement of the first kind (e.g. Busch *et al* [5]).

In the literature on the foundations of quantum mechanics, measurements of the first kind are very popular, presumably because the simple schemes (1), (2) and (4) allow us to make deductions in a relatively simple manner. However, it was realized by Wigner [6] that these schemes cannot be generally valid, in particular if there is an additive constant of the motion. Also the problems with observables having a continuous spectrum are well known. These problems have been a reason for doubting the general validity of the von Neumann–Lüders projection as well as of the concept of measurement of the first kind (see e.g. [7]). Doubts regarding their necessity as characteristics of quantum-mechanical measurement procedures are also generated by the circumstance that widely used measurement procedures do not satisfy them. Thus, in an ideal photon counter the photons are annihilated, ideally yielding the vacuum state as the final state of the electromagnetic field for any initial state. Minimal disturbance of photon number would be possible in this measurement process only on the basis of a certain inefficiency of the detection process.

Also the Stern–Gerlach measurement procedure for measuring a spin component, being a paradigm of first-kind measurements, satisfies this characteristic only in an approximate sense [8, 9]. So that the Stern–Gerlach device works properly, it is necessary that the magnetic field \mathbf{B} is inhomogeneous. Because the magnetic field must satisfy Maxwell's equations, and, hence, should satisfy the requirement $\nabla \cdot \mathbf{B} = 0$, such an inhomogeneity is inconsistent with the assumption generally made in textbooks of quantum mechanics, that the direction of \mathbf{B} is the same everywhere. Hence, the interaction Hamiltonian $-(\mu/2)\mathbf{B} \cdot \boldsymbol{\sigma}$ does not contain just one single spin component but contains different non-commuting Pauli spin matrices, thus preventing a single spin component from being conserved. Since the inhomogeneity of the magnetic field is essential to the functioning of the Stern–Gerlach measurement, its deviation from first-kindness is even crucial.

The examples discussed above suggest that the concept of a measurement of the first kind at most has a theoretical importance, and is seldom realized in practice. It is important to note that the concept, being a requirement to be met by the post-measurement state of the object, hinges strongly on the *preparative* aspect of measurement. When we take the point of view that measurements are performed in order to obtain knowledge about the state of the object immediately *preceding* the measurement, then it is not very important in which state the object is left after the measurement. From this *determinative* point of view it is completely unimportant whether a measurement procedure satisfies the von Neumann–Lüders prescription or not. In particular, we do not have any reason to require the final object state to be prepared according to the projection (4).

In general, the final states of both object and measuring instrument are determined by the measurement interaction. Let ρ_a be the density operator of the initial state of the measuring instrument, and $U = e^{-\frac{i}{\hbar}HT}$. Then $\rho_f = U\rho_a U^\dagger$ is the final state of the combined system. From a determinative point of view the final state ρ_{af} of the measuring instrument is the crucial quantity because it yields the probabilities of the pointer positions that are directly compared with the experimental data:

$$p_m = \text{Tr}_{\mathcal{H}_a} \rho_{af} E_{am} \quad \rho_{af} = \text{Tr}_{\mathcal{H}_o} \rho_f \quad (6)$$

in which the operators E_{am} define the spectral representation of the pointer observable. These probabilities can be written according to

$$p_m = \text{Tr}_{\mathcal{H}_o} \rho \text{Tr}_{\mathcal{H}_a} \rho_a U^\dagger E_{am} U. \quad (7)$$

Defining $M_m := \text{Tr}_{\mathcal{H}_a} \rho_a U^\dagger E_{am} U$ the probabilities (7) can be represented as

$$p_m = \text{Tr} \rho M_m \quad M_m = M_m^\dagger \geq 0 \quad \sum_m M_m = I \quad (8)$$

thus enabling us to consider them as properties of the initial object state ρ .

It is important to note that the operators M_m in (8) in general are not projection operators, as would have been obtained if the probabilities were related to the spectral representation of a Hermitian operator, as is the case in the Dirac–von Neumann formulation of quantum mechanics. By now it is well known (e.g. [10, 11, 3, 5]), however, that the latter formalism is too restrictive to encompass all experimental measurement procedures used in actual practice. The probability distributions of most quantum measurements should be given by (8) in which the set of operators $\{M_m\}$ constitute a decomposition of the identity generating a positive operator-valued measure (POVM) (for applications to photon detector and Stern–Gerlach, see e.g. [12, 9]). Measurements within the domain of application of the Dirac–von Neumann formalism just correspond to a subclass described by *orthogonal* decompositions of the identity generating projection-valued measures (PVM). A second reason to doubt the general validity of the von Neumann–Lüders projection is that its formulation as given above only applies to measurements represented by the Hermitian operators of the Dirac–von Neumann formalism.

The final object state ρ_{of} is obtained from ρ_f by partial tracing over the apparatus Hilbert space: $\rho_{of} = \text{Tr}_{\mathcal{H}_a} \rho_f$. This state is completely determined by the initial conditions and the measurement interaction. For each measurement procedure corresponding to a PVM it is possible to verify whether this state equals (3). For most realistic measurement procedures this is not the case (see [13] for a discussion of a widely used model for measuring PVMs). It is straightforward to see that for arbitrary measurements of POVMs the state ρ_{of} can be given according to

$$\rho_{of} = \sum_m p_m \rho_{ofm} \quad \rho_{ofm} = \frac{\text{Tr}_{\mathcal{H}_a} E_{am} \rho_f}{p_m}. \quad (9)$$

Here the density operator ρ_{ofm} describes the object’s post-measurement state conditional on the measurement result m [14]. Only for very particular measurement interactions does this coincide with the state $E_m \rho E_m / \text{Tr} \rho E_m$ prescribed by the von Neumann–Lüders projection. By Busch *et al* [5] it is proposed to generalize the definition of first-kind measurements to measurements described by a POVM, so as to satisfy an equality analogous to (5),

$$\text{Tr} \rho_{of} M_m = \text{Tr} \rho M_m \quad (10)$$

in which ρ_{of} is the post-measurement object state (9). However, since this equality hinges on the preparative aspect of measurement, as before it seems to impose on the interaction

between object and measuring instrument requirements that are both unnecessary and hard to meet in actual practice.

It is the purpose of this paper to propose an alternative to the Lüders projection (4), generalizing the von Neumann projection to measurements described by POVMs, that is both relevant in a *determinative* sense and valid for all such measurements. This is done in section 2. In section 3 a number of applications of the alternative projection will be discussed. Finally in section 4 an interpretation of the projected state will be given deviating from the interpretation as a description of the final object state.

2. Generalized von Neumann projection

In this section a generalization of the von Neumann projection (2) is defined, valid for measurements represented by arbitrary POVMs $\{M_m\}$ with linearly independent operators M_m . For simplicity we shall restrict the general formulation to finite-dimensional Hilbert spaces, although an extension to infinite dimensions seems to be straightforward. Then the operators M_m can be considered as vectors in Hilbert–Schmidt space \mathcal{H}_{HS} , the set $\{M_m\}$ constituting a (generally non-orthogonal) basis of a subspace $\mathcal{H}_{\{M_m\}}$. Within this subspace a bi-orthonormal system (cf appendix A) is constructed by requiring

$$\begin{aligned} M'_{m'} &= \sum_m \beta_{m'm} M_m \\ \langle M_m, M'_{m'} \rangle &= \text{Tr } M_m M'_{m'} = \delta_{mm'}. \end{aligned} \quad (11)$$

Since the coefficients $\beta_{m'm}$ are real we have $M'_m = M_m^\dagger$.

Within Hilbert–Schmidt space we now define the non-orthogonal projections (cf appendix A) $|M_m\rangle\langle M'_m|$ and $|M'_m\rangle\langle M_m|$, which are in general non-Hermitian (super)operators. However, analogously to (22) the projection operator

$$\mathcal{P}_{\{M_m\}} = \sum_m |M_m\rangle\langle M'_m| = \sum_m |M'_m\rangle\langle M_m| \quad (12)$$

(summation over all elements of the decomposition of the identity) is a Hermitian operator on Hilbert–Schmidt space, $\mathcal{P}_{\{M_m\}} = \mathcal{P}_{\{M_m\}}^*$. Due to hermiticity it equals the unique orthogonal projection onto the subspace $\mathcal{H}_{\{M_m\}}$,

$$\mathcal{P}_{\{M_m\}} \mathcal{H}_{\text{HS}} = \mathcal{H}_{\{M_m\}}.$$

Using (12), the generalized von Neumann projection is defined according to

$$\rho \rightarrow \mathcal{P}_{\{M_m\}}(\rho) = \sum_m (\text{Tr } M'_m \rho) M_m = \sum_m (\text{Tr } M_m \rho) M'_m. \quad (13)$$

It is easily seen that $\mathcal{P}_{\{M_m\}}(\rho)$ reduces to (3) if the POVM is a maximal PVM for which the operators M_m are one-dimensional projection operators. From (11) it also immediately follows that

$$\text{Tr } \mathcal{P}_{\{M_m\}}(\rho) M_m = \text{Tr } \rho M_m. \quad (14)$$

In contrast to (10), this equality holds true for any quantum-mechanical measurement procedure represented by a POVM.

The mapping (13) is a mi-morphism. This is proven as follows.

(i)

$$\mathcal{P}_{\{M_m\}} \left(\sum_m (\text{Tr } M_m \rho) M'_m \right) = \sum_m (\text{Tr } M_m \rho) \sum_{m'} (\text{Tr } M_{m'} M'_m) M'_{m'}$$

yielding $\sum_m (\text{Tr } M_m \rho) M'_m$ because of (11). Hence, $\mathcal{P}_{\{M_m\}}^2 = \mathcal{P}_{\{M_m\}}$.

(ii)

$$\left. \begin{aligned} \text{Tr } M_m M'_{m'} &= \delta_{mm'} \\ \sum_m M_m &= I \end{aligned} \right\} \rightarrow \text{Tr } M'_m = 1.$$

Hence

$$\text{Tr } \mathcal{P}_{\{M_m\}}(\rho) = \text{Tr} \left(\sum_m (\text{Tr } M_m \rho) M'_m \right) = \text{Tr } \rho.$$

(iii) The projector $\mathcal{P}_{\{M_m\}}$ onto the subspace $\mathcal{H}_{\{M_m\}}$ is a Hermitian (super)operator. Because the operators M'_m are in general not non-negative, the property $\mathcal{P}_{\{M_m\}}(\rho) \geq O$ cannot easily be proven using definition (13). However, because of Naimark's theorem (cf appendix B) there exists a second Hermitian projector onto the subspace $\mathcal{H}_{\{M_m\}}$, namely

$$\mathcal{P}_{\{M_m\}}^{\text{orth}}(A) = P A P$$

$P = P^\dagger$ the Hermitian projection operator constructed by the Naimark theorem by relating the operators M_m to a PVM $\{E_m\}$ according to $M_m = P E_m P$; hence $M_m = P M_m P$ and

$$\mathcal{P}_{\{M_m\}}^{\text{orth}} \mathcal{H}_{\text{HS}} = \mathcal{H}_{\{M_m\}}.$$

Because of the uniqueness of Hermitian projectors we must have

$$\mathcal{P}_{\{M_m\}} = \mathcal{P}_{\{M_m\}}^{\text{orth}}.$$

Then

$$P \rho P > O \rightarrow \mathcal{P}_{\{M_m\}}(\rho) > O.$$

3. Applications

3.1. Non-maximal PVMs

Let $\{E_m\}$ be a non-maximal PVM, with $N_m = \text{Tr } E_m$ the dimension of subspace $E_m \mathcal{H}$. Then

$$\text{Tr } E_m E'_{m'} = \delta_{mm'} \rightarrow E'_m = E_m / N_m.$$

Hence

$$\mathcal{P}_{\{E_m\}}(\rho) = \sum_m (\text{Tr } \rho E_m) E_m / N_m.$$

For maximal PVMs (for which $N_m = 1 \forall m$) this equals result (3) of the von Neumann projection. In general it is, evidently, different from the result of the Lüders projection (4).

3.2. Invertible non-ideal measurement of PVM $\{E_m\}$

In [12] a non-ideal measurement of the PVM $\{E_m\}$ was defined as a measurement represented by the POVM $\{M_m\}$ satisfying

$$M_m = \sum_{m'} \lambda_{mm'} E_{m'} \quad \lambda_{mm'} \geq 0 \quad \sum_m \lambda_{mm'} = 1.$$

A non-ideal measurement is called invertible if the non-ideality matrix $(\lambda_{mm'})$ has a unique inverse $(\lambda_{m'm}^{-1})$. Then

$$M'_m = \sum_{m'} \frac{\lambda_{m'm}^{-1}}{N_{m'}} E_{m'}$$

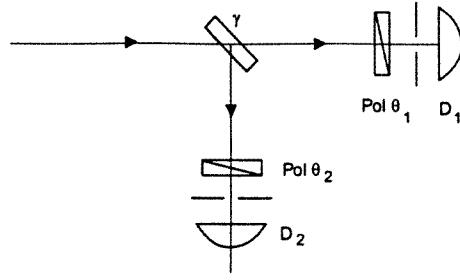


Figure 1. Joint non-ideal measurement of two polarization observables.

and

$$\mathcal{P}_{\{M_m\}}(\rho) = \sum_m (\text{Tr } E_m \rho) E_m / N_m.$$

It is important to note that this result equals the one obtained for ideal measurements of the PVM $\{E_m\}$ as obtained in section 3.1, both for maximal and non-maximal PVMs.

3.3. Joint non-ideal measurement of incompatible PVMs $\{E_m\}$ and $\{F_n\}$

As a third application we consider the joint non-ideal measurement of incompatible observables, defined in [12] according to the requirement that the measurement is represented by a bivariate POVM $\{M_{mn}\}$ satisfying

$$\begin{aligned} \sum_n M_{mn} &= \sum_{m'} \lambda_{mm'} E_{m'} & \lambda_{mm'} &\geq 0 & \sum_m \lambda_{mm'} &= 1 \\ \sum_m M_{mn} &= \sum_{n'} \mu_{nn'} F_{n'} & \mu_{nn'} &\geq 0 & \sum_n \mu_{nn'} &= 1. \end{aligned}$$

As an example we shall calculate the projected density operator for the joint non-ideal measurement of two polarization observables by the measurement arrangement of figure 1, in which a semitransparent mirror (transmission coefficient γ , $0 < \gamma < 1$) has outgoing beams into the directions of polarizing nicols oriented in different directions θ_1 and θ_2 , respectively.

For the present purpose it is unimportant that this measurement can be interpreted as a joint measurement. Hence, we can ignore the possibility to write the POVM in a bivariate form. Then the POVM is given as $\{M_1 = \gamma E_+, M_2 = (1-\gamma)F_+, M_3 = \gamma E_- + (1-\gamma)F_-\}$, representing the probabilities that the photon is detected either in detector D_1 or in D_2 , or is absorbed in one of the nicols. Using the representation generated by the first nicol the operators can be given according to

$$\begin{aligned} E_+ &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & F_+ &= \begin{pmatrix} p & \sqrt{p(1-p)} \\ \sqrt{p(1-p)} & 1-p \end{pmatrix} \\ E_- &= I - E_+ & F_- &= I - F_+ \\ p &= \text{Tr } E_+ F_+ = \cos^2(\theta_1 - \theta_2). \end{aligned}$$

From (11) we obtain

$$\begin{cases} M'_1 = \{2\gamma p(1-p)\}^{-1}[\{1-\gamma(1-p)\}E_+ \\ \quad + \{(1-\gamma)(1-p) - p\}F_+ + (1-p)\{\gamma p - (1-\gamma)\}I] \\ M'_2 = \{2(1-\gamma)p(1-p)\}^{-1}[\{\gamma(1-p) - p\}E_+ \\ \quad + \{1 - (1-\gamma)(1-p)\}F_+ + (1-p)\{(1-\gamma)p - \gamma\}I] \\ M'_3 = \{2p\}^{-1}[-E_+ - F_+ + (1-p)I] \end{cases}$$

yielding

$$\mathcal{P}_{\{M_m\}}(\rho) = \frac{1}{2p(1-p)} \begin{cases} [p(1-p+p_1-p_2)E_+ + p(1-p+p_2-p_1)F_+ \\ \quad + (1-p)(1+p-p_1-p_2)E_- \\ \quad + (1-p)(1+p-p_1-p_2)F_-] \end{cases} \quad (15)$$

$$p_1 = \text{Tr } \rho E_+ \quad p_2 = \text{Tr } \rho F_+.$$

Although from the general theory it follows that this operator is non-negative, this is not easily seen from its representation as given here, because for certain values of p , p_1 and p_2 the coefficients may be negative (for instance, $1-p+p_1-p_2 < 0$ if $p = \frac{9}{10}$, $p_1 = \frac{1}{4}$, $p_2 = \frac{1}{2}$). A direct proof of the non-negativity of (15) can be given, though, taking into account that p , p_1 and p_2 should satisfy the triangle inequality

$$(2(1-p))^{1/2} \leq (\text{Tr } \rho^2 + 1 - 2p_1)^{1/2} + (\text{Tr } \rho^2 + 1 - 2p_2)^{1/2}.$$

The possibility of negative coefficients in (15) has important consequences for the interpretation of the state function $\mathcal{P}_{\{M_m\}}(\rho)$: it is evidently not possible to interpret this density operator as representing a mixture of states having sharp values of either $\{E_m\}$ or $\{F_n\}$.

3.4. Eight-port optical homodyning

As a final example we consider eight-port optical homodyning (see figure 2), which is a detection method for monochromatic optical signals, known to be represented by the POVM $M(q, p) = |\alpha\rangle\langle\alpha|/2\pi$, $\alpha = (1/\sqrt{2})(q + ip)$, $|\alpha\rangle$ a coherent state [15]. This example is actually an infinite-dimensional one. Moreover, it is complicated by the continuity of the variables $-\infty < q < \infty$, $-\infty < p < \infty$. Nevertheless, the methods developed before for

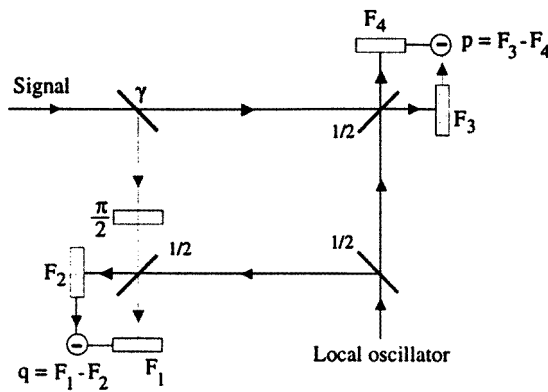


Figure 2. Eight-port optical homodyning.

the finite-dimensional case will be seen to work properly also here if the mathematics is treated in the usual sense proposed by Dirac, thus demonstrating that the extension of the theory to infinite-dimensional Hilbert spaces does not seem to meet essential difficulties.

Generalizing (11) according to

$$\text{Tr } M(q, p) M'(q', p') = \delta(q - q') \delta(p - p')$$

it is straightforward to prove that the operators $M'(q, p)$ can be represented according to

$$M'(q, p) = \frac{1}{2\pi} \int du dv e^{\frac{u^2+v^2}{4}} e^{i(uq+vp)} e^{-i(uQ+vP)} = M'(q, p)^\dagger. \quad (16)$$

Then the projected density operator is given according to

$$\mathcal{P}_{\{M(q,p)\}}(\rho) = \int dq dp \left(\text{Tr} \frac{|\alpha\rangle\langle\alpha|}{2\pi} \rho \right) M'(q, p). \quad (17)$$

It is straightforward to prove that

$$\mathcal{P}_{\{M(q,p)\}}(\rho) = \rho. \quad (18)$$

Result (18) is an important one, demonstrating that eight-port optical homodyning is a *complete* measurement allowing us (at least in principle) to calculate the density operator ρ of the initial state from the probability distribution $\text{Tr}(|\alpha\rangle\langle\alpha|/2\pi)\rho$ obtained in the measurement. This result can also be obtained by employing the fact that the eight-port homodyning measurement can be interpreted as a joint non-ideal measurement of the observables $Q = (a + a^\dagger)/\sqrt{2}$ and $P = -i(a - a^\dagger)/\sqrt{2}$ of the monochromatic electromagnetic mode [16].

4. Interpretation

4.1. Mathematical interpretation

Results (17) and (18) exhibit an important link between the projected state function $\mathcal{P}_{\{M_m\}}(\rho)$ and phase-space representations of quantum mechanics [17, 18], the most well known being the Wigner–Weyl representation

$$\begin{aligned} \rho_{\text{ww}}(q, p) &= \text{Tr } \rho W(q, p) \\ W(q, p) &= \frac{1}{(2\pi)^2} \int du dv e^{i(uq+vp)} e^{-i(uQ+vP)} = W(q, p)^\dagger \end{aligned}$$

where Q and P are position and momentum operators as given above. The operators $W(q, p)$ can be considered as constituting (in the Dirac sense) a complete orthogonal set of vectors in Hilbert–Schmidt space, since

$$\text{Tr } W(q, p) W(q', p') = \frac{1}{2\pi} \delta(q - q') \delta(p - p')$$

and

$$A = 2\pi \int dq dp (\text{Tr } A W(q, p)) W(q, p) \quad A \in \mathcal{H}_{\text{HS}}.$$

Defining

$$W'(q, p) = 2\pi W(q, p)$$

we see that the sets $\{W(q, p)\}$ and $\{W'(q, p)\}$ constitute a bi-orthonormal system in \mathcal{H}_{HS} . Hence, the Wigner–Weyl representation corresponds to an expansion of Hilbert–Schmidt vectors using the orthogonal basis $\{W(q, p)\}$.

Analogously, the so-called Ω -representations can be interpreted as expansions using non-orthogonal bases for Hilbert–Schmidt space. Thus,

$$\begin{aligned} \rho &= \int dq dp (\text{Tr } \rho W_\Omega(q, p)^\dagger) W'_\Omega(q, p) \\ W_\Omega(q, p) &= \frac{1}{(2\pi)^2} \int du dv e^{i(uq+vp)} \Omega(u, v) e^{-i(uQ+vP)} \\ W'_\Omega(q, p) &= \frac{1}{2\pi} \int du dv e^{-i(uq+vp)} \Omega(u, v)^{-1} e^{i(uQ+vP)} \\ \text{Tr } W_\Omega(q, p)^\dagger W'_\Omega(q', p') &= \delta(q - q') \delta(p - p'). \end{aligned}$$

Note that the Wigner–Weyl representation is obtained for $\Omega(q, p) = 1$. With $\Omega(q, p) = \exp\{-(u^2 + v^2)/4\}$ we have a special case of the Husimi representation [19], yielding $W_\Omega(q, p) = |\alpha\rangle\langle\alpha|/2\pi$ and $W'_\Omega(q, p) = M'(q, p)$ as given by (16). Hence, the elements of the POVM of eight-port homodyning constitute (in the Dirac sense) a complete non-orthogonal set of vectors in Hilbert–Schmidt space, this being the mathematical source of the completeness of this measurement procedure.

It is now possible to see how things work out if a measurement is not a complete one. In that case the elements M_m do not span the whole Hilbert–Schmidt space, but only a subspace $\mathcal{H}_{\{M_m\}}$. For this reason the measurement does not yield complete information about the initial state ρ . Actually, a measurement of an incomplete POVM $\{M_m\}$ only allows us to obtain information on the state $\mathcal{P}_{\{M_m\}}(\rho)$, which is just the projection of ρ onto the subspace $\mathcal{H}_{\{M_m\}}$. If $\rho \in \mathcal{H}_{\{M_m\}}$ then $\mathcal{P}_{\{M_m\}}(\rho) = \rho$, and the measurement yields complete information on the initial state. However, this measurement does not yield any information on the part of ρ that is in subspace $(I - \mathcal{P}_{\{M_m\}})\mathcal{H}_{\text{HS}}$. Evidently, the density operator $\mathcal{P}_{\{M_m\}}(\rho)$ represents the information regarding the initial state ρ , to be obtained by the measurement of the POVM $\{M_m\}$.

From this perspective it is clear that measurements represented by the Hermitian operators of the Dirac–von Neumann formalism (i.e. PVMs) are not singled out in any special way within the class of all possible measurement procedures. Comparing applications 3.1 and 3.2 we see that the projected density operators are the same for an ideal and non-ideal measurement of a PVM if the latter is an invertible one. We shall interpret this as *informational equivalence* of these measurements, caused by the fact that the operators of the POVM $\{M_m\}$ span the same subspace of Hilbert–Schmidt space as the operators of the PVM $\{E_m\}$. However, as is now becoming abundantly clear [20, 21, 15] many measurement procedures exist yielding more information on ρ because the elements of their POVMs span larger subspaces. Applications 3.3 and 3.4 are examples of such measurements. In principle the subspace of Hilbert–Schmidt space spanned can range from a one-dimensional one (in case of an uninformative measurement represented by the POVM consisting only of the operator I) to the whole of Hilbert–Schmidt space (in case of a complete measurement). It seems that the emphasis in the literature on the foundations of quantum mechanics on measurements represented by PVMs can hardly be justified by the experimental importance of such measurements (for instance, all scattering experiments determining differential cross sections should be analysed in terms of (positive) operator-valued measures).

Since only the part $\mathcal{P}_{\{M_m\}}(\rho)$ of ρ that is in $\mathcal{H}_{\{M_m\}}$ can be reconstructed from the data obtained by the measurement of the POVM $\{M_m\}$ there is an information deficit associated with any incomplete measurement, preventing a complete reconstruction of ρ . As a quantitative measure of the information deficit it is possible to take the quantity $D_{\{M_m\}}(\rho) := H(\mathcal{P}_{\{M_m\}}(\rho)) - H(\rho)$, in which $H(\rho)$ is the von Neumann entropy defined by

$H(\rho) = -\text{Tr } \rho \ln \rho$. It is possible to prove (see appendix C) that for arbitrary POVM $\{M_m\}$

$$D_{\{M_m\}}(\rho) \geq 0. \quad (19)$$

Evidently, for complete measurements $D_{\{M_m\}}(\rho) = 0$. For the uninformative measurement represented by the POVM $\{I\}$ we find $\mathcal{P}_{\{I\}}(\rho) = I/N$, yielding $H(\mathcal{P}_{\{I\}}(\rho)) = \ln N$, and $D_{\{I\}}(\rho) = \ln N - H(\rho)$. Hence $D_{\{I\}}(\rho)$ is reaching its maximal value $\ln N$ for pure states ρ , the information deficit being smaller if the initial state is a mixture.

4.2. Physical interpretation

As already noted before we do not have any reason to suppose that the projected state $\mathcal{P}_{\{M_m\}}(\rho)$ has any simple relation to the final object state ρ_{of} of the measurement. Although for the Lüders projection (4) we have $H(\rho_{of}) - H(\rho) \geq 0$ (which is often thought to be an attractive property because the measurement process is thought to increase entropy), the example of the ideal photon counter (ideally satisfying $H(\rho_{of}) = 0$) demonstrates that measurement disturbance need not imply that the distribution of the final state over the possible output channels is more disorderly than the distribution of the initial state over the input channels. Whereas the Lüders projection is a fruit of a ‘minimal disturbance’ philosophy yielding entropy increase, it is seen from the example that a larger disturbance may increase order in the sense of concentrating the output in a single channel. This is a different way of disqualifying the Lüders projection as an attractive property of quantum measurement.

Since no entropy increase is required (in the sense defined by the von Neumann entropy) in the transition from initial to final object state of a quantum-mechanical measurement, the generalized projection (13), satisfying (19), must have a different interpretation. As found in section 4.1, the projected state $\mathcal{P}_{\{M_m\}}(\rho)$ contains information on the *initial* rather than the final object state of the measurement. For this reason it seems reasonable not to associate (13) with a transition from initial to final state, but interpret it as a transition between *different descriptions of the initial state*. Then $\mathcal{P}_{\{M_m\}}(\rho)$ might be seen as a description of the initial state as far as ‘observable’ by means of a measurement of the POVM $\{M_m\}$.

An interpretation in the sense proposed here is a rather natural one, and not specific to quantum mechanics. As a matter of fact, it is widely used in many domains of physics. Thus, for instance a rigid-body description of a billiard ball is valid if the description is restricted to those observations that are insensitive to the internal dynamics of the constituting atoms. Within the context of such measurements the rigid-body model yields a description that is comparable with the description by the state function $\mathcal{P}_{\{M_m\}}(\rho)$. A description by means of the density operator ρ is necessary if information also has to be dealt with that is valid within the contexts of measurements of POVMs different from $\{M_m\}$.

Continuing the billiard ball analogy, it might be asked whether the state $\mathcal{P}_{\{M_m\}}(\rho)$ could, maybe, be interpreted in a contextualistic-realist sense as the state of the object as it *is* within the context of the measurement arrangement of the POVM $\{M_m\}$. Thus, a billiard ball *is* a rigid body as long as the experimental context allows it to ‘be’ one (and not just ‘appears to be one’). Such a contextualistic-realist notion is sometimes implied by the Copenhagen interpretation, although in this latter interpretation the contextual state is often associated with the final rather than with the initial object state. As suggested by the foregoing, if $\mathcal{P}_{\{M_m\}}(\rho)$ would describe such a *contextual* state, then it should refer to the initial rather than to the final state of the measurement. This would mean that the quantum-mechanical reality corresponding to it must be shaped by the presence of the measurement arrangement even before the interaction between object and measuring instrument has started. Thus, the

von Neumann projection (1) might be thought to be interpretable as the process in which the value of the measured (Dirac–von Neumann) observable comes into being at the first contact of object and measuring instrument, the subsequent measurement process serving to translate this microscopic event into a directly observable macroscopic one.

Because, due to (14), the states ρ and $\mathcal{P}_{\{M_m\}}(\rho)$ yield the same quantum-mechanical measurement result for the POVM $\{M_m\}$ we do not have any means to test whether the initial state within the context of the measurement is really one or the other. For this reason the contextualistic-realist interpretation can neither be verified nor falsified experimentally. However, the examples of generalized measurements discussed above do not seem to favour such an interpretation. Thus, in the case of the inefficient photon detector it is much more natural to interpret the measurement result n as the number of photons registered by the detector (out of the number $m > n$ that were actually present), then as the result of an ideal registration of a non-ideal number of photons being present in the initial state. Also, example 3.3 is not particularly favourable to an interpretation of the contextual state (15) as a state in which either one or both of the observables measured jointly (be it non-ideally) would have adopted a certain value of the measured observable within the context of the measurement arrangement. For these reasons a contextualistic-realist interpretation of the state $\mathcal{P}_{\{M_m\}}(\rho)$ does not seem possible surpassing an attribution of a more or less symbolic meaning as representing a preparation procedure in which, apart from the preparing apparatus, also the measuring instrument (i.e. Bohr’s *whole* measurement arrangement) is taken into account. It seems that such an interpretation of the state $\mathcal{P}_{\{M_m\}}(\rho)$ could be accommodated in a natural way in the empiricist interpretation of quantum mechanics developed in [22].

Acknowledgments

I thank professor J J Seidel for drawing to my attention to his proof of the lemma given in appendix B. I also thank Jan Hilgevoord for a valuable remark.

Appendix A. Non-orthogonal projections in Hilbert space and bi-orthonormal systems

Let $\{e_1, \dots, e_N\}$ be a non-orthogonal basis in an N -dimensional Hilbert space \mathcal{H} . A bi-orthonormal system is defined in the following way:

$$\langle f_n, e_m \rangle = \delta_{nm}. \tag{20}$$

Then also the vectors $\{f_1, \dots, f_N\}$ constitute a non-orthogonal basis for \mathcal{H} .

Using Dirac notation we define the operators

$$\mathcal{P}_n := |e_n\rangle\langle f_n| \tag{21}$$

satisfying

$$\mathcal{P}_n^2 = \mathcal{P}_n \quad \mathcal{P}_n \mathcal{P}_m = O \quad n \neq m.$$

Hence the operator \mathcal{P}_n is a projection operator projecting onto e_n parallel to the vectors $e_{n'}$ with $n' \neq n$, and, hence, in general non-orthogonally. From

$$\mathcal{P}_n^* = |f_n\rangle\langle e_n| \neq \mathcal{P}_n$$

it is evident that the projection operators are non-Hermitian unless the basis $\{e_1, \dots, e_N\}$ is an orthonormal one (in which case $f_n = e_n$, and the projections (21) are orthogonal ones).

Let $\mathcal{H}_{\{e_1, \dots, e_M\}}$ be the subspace spanned by the vectors $\{e_1, \dots, e_M\}$, $M \leq N$. Then the operator $\mathcal{P}_{\{M\}} = \sum_{n=1}^M \mathcal{P}_n$ is a non-orthogonal projection onto this subspace, i.e.

$\mathcal{P}_{\{M\}}\mathcal{H} = \mathcal{H}_{\{e_1, \dots, e_M\}}$. Analogously we have $\mathcal{P}_{\{M\}}^* = \sum_{n=1}^M \mathcal{P}_n^*$, and $\mathcal{P}_{\{M\}}^*\mathcal{H} = \mathcal{H}_{\{f_1, \dots, f_M\}}$, demonstrating that $\mathcal{P}_{\{M\}}$ and $\mathcal{P}_{\{M\}}^*$ have different subspaces as their ranges, unless we start from an orthonormal basis. However, for any bi-orthonormal system we have

$$\sum_{n=1}^N \mathcal{P}_n = \sum_{n=1}^N \mathcal{P}_n^* = I. \quad (22)$$

To a subspace of \mathcal{H} many projections correspond, having this subspace as its range. However, among these there is only one unique projection operator that is Hermitian. This is the projection operator corresponding to the orthogonal projection onto the subspace.

Appendix B. Proof of Naimark's theorem for maximal POVMs

By Naimark [23, 11, 24] the following theorem has been proven.

Naimark's theorem. If $\{M_m\}$ is a POVM on a Hilbert space \mathcal{H} , then there exists a Hilbert space \mathcal{H}' , a PVM $\{E'_m\}$ on \mathcal{H}' , and an orthogonal projection operator P , such that $P\mathcal{H}' = \mathcal{H}$ and $M_m = PE'_mP$.

In this appendix an elementary proof of the theorem is given for finite-dimensional \mathcal{H} (dimension = N), and a POVM $\{M_m\}$ generated by a non-orthogonal decomposition of the identity consisting of M ($M > N$) elements, all being, up to a positive multiplicative constant, equal to a one-dimensional projection operator (so-called maximal POVM [12]). Thus

$$M_m = \alpha_m P_m \quad \alpha_m \geq 0 \quad P_m = |\psi_m\rangle\langle\psi_m| = P_m^2.$$

In order to prove the theorem we first prove a lemma on so-called *eutactic* stars (cf Coxeter [25]), consisting of an over-complete set of vectors $\{e_1, \dots, e_M\}$ spanning an N -dimensional Hilbert space ($M > N$), that can be obtained by orthogonal projection from an orthonormal basis in an M -dimensional Hilbert space \mathcal{H}' .

Lemma (cf Seidel [26]). A star is eutactic if and only if its Gram matrix has precisely two eigenvalues (one of which equals 0).

Proof.

(i) Suppose

$$\begin{aligned} e_i &= Pc_i & i &= 1, \dots, M \\ \langle c_i | c_j \rangle &= \delta_{ij} & i, j &= 1, \dots, M \\ P &= P^\dagger = P^2. \end{aligned}$$

Then

$$\langle e_i | e_j \rangle = \langle c_i | Pc_j \rangle$$

implying that the Gram matrix is a matrix representation of an orthogonal projection operator P . Hence the (Hermitian) Gram matrix has only eigenvalues 0 and 1.

(ii) Conversely, suppose that the $M \times M$ matrix $(\langle e_i | e_j \rangle)$ has only two eigenvalues. At least one of these must be 0, since, because of the linear dependence of the vectors e_i , $i = 1, \dots, M$, we have $\text{Det}(\langle e_i | e_j \rangle) = 0$. We also have $\text{Det}(\langle e_i | e_j \rangle) = \prod_k \lambda_k$, with λ_k the eigenvalues of the Gram matrix. Because the Gram matrix has rank N there are exactly $M - N$ eigenvalues 0.

Given that there is only one non-vanishing eigenvalue, it is possible by means of normalization of the vectors e_i to set this eigenvalue equal to 1. This implies that a unitary $M \times M$ matrix (U_{ij}) exists, such that

$$\sum_{j,k=1}^M U_{ij} \langle e_j | e_k \rangle U_{k\ell}^\dagger = P_{i\ell} \tag{23}$$

$$P_{ii} = 1 \quad i = 1, \dots, N \quad P_{i\ell} = 0 \quad \text{for all other values of } i \text{ and } \ell.$$

Define

$$\mathbf{q}_i = \sum_{j=1}^M U_{ij}^* e_j \quad i = 1, \dots, M. \tag{24}$$

Then (23) can be written as

$$\langle \mathbf{q}_i | \mathbf{q}_\ell \rangle = P_{i\ell}.$$

From this it follows that

$$\mathbf{q}_i = \mathbf{0} \quad i = N + 1, \dots, M$$

and

$$e_j = \sum_{i=1}^M (U^{-1})_{ji}^* \mathbf{q}_i = \sum_{i=1}^N (U^{-1})_{ji}^* \mathbf{q}_i.$$

Moreover, the vectors $\mathbf{q}_i, i = 1, \dots, N$ constitute an orthonormal set. Because of (24) this is an orthonormal basis of \mathcal{H} .

Taking \mathcal{H} as a subspace of an M -dimensional Hilbert space \mathcal{H}' we obtain an orthonormal basis in this latter space by supplementing the set $\mathbf{q}_i, i = 1, \dots, N$ with $M - N$ mutually orthogonal normalized vectors in the orthogonal complement of \mathcal{H} in \mathcal{H}' . The orthonormal basis of \mathcal{H}' thus obtained is indicated as $\tilde{\mathbf{q}}_i, i = 1, \dots, M$ (hence, $\tilde{\mathbf{q}}_i = \mathbf{q}_i, i = 1, \dots, N$). When P is the orthogonal projection operator within \mathcal{H}' onto the subspace \mathcal{H} , then for this basis:

$$\begin{aligned} P \tilde{\mathbf{q}}_i &= \mathbf{q}_i & i = 1, \dots, N \\ P \tilde{\mathbf{q}}_i &= \mathbf{0} & i = N + 1, \dots, M. \end{aligned} \tag{25}$$

A different orthonormal basis is construed in \mathcal{H}' by means of the unitary transformation

$$\mathbf{c}_j = \sum_{i=1}^M (U^{-1})_{ji}^* \tilde{\mathbf{q}}_i \quad j = 1, \dots, M.$$

Using (25) this finally implies

$$P \mathbf{c}_j = P \sum_{i=1}^M (U^{-1})_{ji}^* \tilde{\mathbf{q}}_i = \sum_{i=1}^N (U^{-1})_{ji}^* \tilde{\mathbf{q}}_i = e_j \quad j = 1, \dots, M.$$

□

Now the Naimark theorem for maximal POVMs follows easily since maximal POVMs can be associated with eutactic stars. Indeed, on \mathcal{H} we have $\sum_{m=1}^M \alpha_m P_m = I$. Defining the (unnormalized) vectors $|\phi_m\rangle = \alpha_m^{1/2} |\psi_m\rangle$ we obtain $\sum_{m=1}^M |\phi_m\rangle \langle \phi_m| = I$, or $\sum_{m=1}^M \langle \phi_k | \phi_m \rangle \langle \phi_m | \phi_\ell \rangle = \langle \phi_k | \phi_\ell \rangle$, implying that the Gram matrix $(\langle \phi_k | \phi_\ell \rangle)$ is a projection. Hence the presuppositions of the lemma are satisfied. Then the projection operators onto the orthonormal basis of \mathcal{H}' construed in the lemma are immediately seen to yield the elements of the PVM sought for. □

Appendix C. Positivity of the information deficit

The proof of (19) uses the well known inequality

$$\mathrm{Tr} A \ln B - \mathrm{Tr} A \ln A \leq \mathrm{Tr} A - \mathrm{Tr} B \quad A, B \geq O \quad (26)$$

that can easily be proven by employing the spectral representations of operators A and B , and by using the convexity of the function $\ln x$, implying the inequality $\ln x \leq 1 - x \forall x$.

Because of the hermiticity of the projection $\mathcal{P}_{\{M_m\}}$ we have

$$\begin{aligned} H(\mathcal{P}_{\{M_m\}}(\rho)) &= -\mathrm{Tr} \mathcal{P}_{\{M_m\}}(\rho) \ln \mathcal{P}_{\{M_m\}}(\rho) \\ &= -\mathrm{Tr} \rho \mathcal{P}_{\{M_m\}}(\ln \mathcal{P}_{\{M_m\}}(\rho)) \\ &= -\mathrm{Tr} \rho \ln e^{\mathcal{P}_{\{M_m\}}(\ln \mathcal{P}_{\{M_m\}}(\rho))}. \end{aligned}$$

Applying inequality (26) with $A = \rho$ and $B = \exp(\mathcal{P}_{\{M_m\}}(\ln \mathcal{P}_{\{M_m\}}(\rho))) = \mathcal{P}_{\{M_m\}}(\rho)$ the inequality now immediately follows because $\mathrm{Tr} A = \mathrm{Tr} B$.

References

- [1] von Neumann J 1932 *Mathematische Grundlagen der Quantenmechanik* (Berlin: Springer)
von Neumann J 1955 *Mathematical Foundations of Quantum Mechanics* (Princeton, NJ: Princeton University Press)
- [2] Lüders G 1951 *Ann. Phys.* **8** 322
- [3] Ludwig G 1983 *Foundations of Quantum Mechanics* vol I and II (Berlin: Springer)
- [4] Furry W H 1966 Quantum theory of measurement *Lectures in Theoretical Physics 8a Statistical Physics and Solid State Physics* ed W E Brittin (Boulder, CO: The University of Colorado Press) pp 1–64
- [5] Busch P, Grabowski M and Lahti P J 1995 *Operational Quantum Mechanics* (Berlin: Springer)
- [6] Wigner E P 1952 *Z. Phys.* **133** 101
- [7] de Muynck W M, Janssen P A E M and Santman A 1979 *Found. Phys.* **9** 71
- [8] Scully M O, Lamb W E Jr and Barut A 1987 *Found. Phys.* **17** 575
- [9] Martens H and de Muynck W M 1993 *J. Phys. A: Math. Gen.* **26** 2001
- [10] Davies E B 1976 *Quantum Theory of Open Systems* (London: Academic)
- [11] Holevo A S 1982 *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North-Holland)
- [12] Martens H and de Muynck W 1990 *Found. Phys.* **20** 255
Martens H and de Muynck W 1990 *Found. Phys.* **20** 357
- [13] Wan K K and Harrison F E 1994 *Found. Phys.* **24** 831
- [14] Ozawa M 1985 *J. Math. Phys.* **26** 1948
- [15] Yamamoto Y, Machida S, Saito S, Imoto N, Yanagawa T, Kitagawa M and Björk G 1990 Quantum mechanical limit in optical precision measurement and communication *Progress in Optics* vol XXVIII, ed E Wolf (Amsterdam: Elsevier) p 87
- [16] de Muynck W M and Martens H 1990 *Coherence and Quantum Optics* vol VI, ed J H Eberly *et al* (New York: Plenum) p 221)
- [17] Agarwal G S and Wolf E 1970 *Phys. Rev. D* **2** 2161
Agarwal G S and Wolf E 1970 *Phys. Rev. D* **2** 2187
Agarwal G S and Wolf E 1970 *Phys. Rev. D* **2** 2206
- [18] Hillery M, O'Connell R F, Scully M O and Wigner E P 1984 *Phys. Rep.* **106** 121
- [19] Husimi K 1940 *Proc. Phys. Math. Soc. Japan* **22** 67
- [20] Busch P 1987 *Found. Phys.* **17** 905
- [21] de Muynck W M and Martens H 1990 *Phys. Rev. A* **42** 5079
- [22] de Muynck W M 1995 *Synthese* **102** 293
- [23] Naimark M 1940 *Izv. Akad. Nauk SSSR Ser. Mat.* **4** 277
- [24] Akhiezer N I and Glazman I M 1981 *Theory of Linear Operators in Hilbert Space* vol II (Boston, MA: Pitman Advanced Publishing Program) p 389
- [25] Coxeter H S M 1973 *Regular Polytopes* (New York: Dover)
- [26] Seidel J J 1976 *Eutactic Stars (Colloquia Mathematica Societas Janos Bolyai 18 Combinatorics)* (Keszthely, Hungary)