PROPERTIES OF THE RANGE OF A STATE OPERATOR

GIANNI CASSINELLI, ERNESTO DE VITO

Department of Physics and I.N.F.N., University of Genova, 16 146 Genova, Italy
(e-mail: cassinelli@genova.infn.it), (e-mail: devito@genova.infn.it)

and

PEKKA J. LAHTI

Department of Physics, University of Turku, SF-20500 Turku, Finland
(e-mail: pekka.lahti@cc.utu.fi)

(Received February 3, 1994)

This paper presents a study of the closure of the range, the range of the square root, and the range of a state operator which arises in the context of quantum theory of measurement as the state of an apparatus after the measurement. The results obtained are placed in the context of the theory of bases in Banach spaces. Physical relevance of these results is also discussed.

1. Introduction

Quantum theory of measurement poses such problems that their solutions have been often looked for outside the quantum mechanics itself. We believe that careful and mathematically precise analysis of the measurement process inside quantum mechanics deserves attention. It provides sharp identification of critical points and blows out the fog of words that often obscures them.

A mathematical model for the measuring process in the standard Hilbert space quantum mechanics has been proposed recently. It has been used both as the source of mathematical investigations and to discuss some interpretational issues in a precisely defined context [1]. Its basic ingredient is a sequence of unit vectors \((\gamma_i)\), in terms of which the coupling between the measured system and the measuring apparatus is described. The state of the apparatus after the measurement is a statistical operator (mixed state), denoted by \(W(\varphi)\). It depends on the initial (vector) state \(\varphi\) of the measured system and on the measurement coupling generating vectors \((\gamma_i)\).

In general, this sequence is completely arbitrary, and the basic problem is to add to the model physically meaningful requirements on the measurements that lead to further specifications of the sequence \((\gamma_i)\). Some of such conditions are already known [1, 2].
The purpose of this paper is to show that some properties of the state operator \( W(\varphi) \), namely the properties of its range, of the range of its square root, and of the closure of its range are equivalent to certain properties of the measurement coupling generating sequence \((\gamma_i)\).

The interest for such characterizations lies in the fact that the properties of the range of the state \( W(\varphi) \) that we are considering have a transparent physical interpretation. The first two refer to the decomposability of the final apparatus state \( W(\varphi) \) in terms of the pointer eigenstates, whereas the third is the starting point for the recently proposed Copenhagen variant of the modal interpretation of quantum mechanics [3]. In fact, various modal interpretations are currently discussed in the literature on the foundations of quantum mechanics. We believe that the results of this paper could constitute the mathematical basis for most of them.

It is our opinion that this paper could be of interest also from the mathematical point of view. In particular, it contains some use of ideas taken from the theory of Schauder bases in Banach spaces. It might be of interest to see these ideas "at work" in Hilbert space.

The framework of this investigation is the ordinary Hilbert space formulation of quantum mechanics whereby the description of a physical system is based on a complex separable Hilbert space \( \mathcal{H} \), with the inner product \( \langle \cdot | \cdot \rangle \). The states and observables of the system are represented as positive trace-one operators \( T \) and as self-adjoint operators \( A \) acting in \( \mathcal{H} \), respectively. If \( P^A \) is the spectral measure of \( A \), then the probability measure defined by this observable and a state \( T \) obtain the explicit form

\[
p^A_T(X) := \text{Tr}[TP^A(X)].
\]  

(1)

According to the minimal (Born) interpretation these numbers are probabilities for measurement outcomes: \( p^A_T(X) \) is the probability that the measurement of observable \( A \) leads to a result in the (Borel) set \( X \) when performed on the system in the state \( T \).

The structure of the paper is as follows. In Sections 2 and 3 we describe the properties of states and the model of measurement on which we base our study. Sections 4, 5, and 6 contain the main results characterizing the closure of the range, the range of the square root, and the range of the state of the apparatus after the measurement in terms of the vectors defining a unitary mapping modelling the measurement coupling. In Section 7 some basic facts of the theory of Schauder bases in Banach spaces are recalled.

2. Properties of states

For an operator \( T \), we let \( \text{ran}(T) \), \( \overline{\text{ran}}(T) \), and \( \ker(T) \) denote its range, the closure of its range, and its kernel, respectively. For any closed subspace \( \mathcal{M} \) of \( \mathcal{H} \) we let \( P[\mathcal{M}] \) denote the projection operator onto \( \mathcal{M} \). If \( \mathcal{M} \) is the one-dimensional subspace generated by a unit vector \( \varphi \), we write \( P[\varphi] \) instead of \( P[\mathcal{M}] \).

2.1. Let \( T(\mathcal{H})^+ \) denote the set of states, that is, bounded positive trace class operators \( T: \mathcal{H} \to \mathcal{H} \) with trace one. It is \( \sigma \)-convex: if \( (T_i) \) is a sequence of states and \( (\lambda_i) \) is
a sequence of weights \( (0 \leq \lambda_i \leq 1, \sum \lambda_i = 1) \), then the series \( \sum \lambda_i T_i \) converges (in the trace norm) to an element of \( T(\mathcal{H})^+ \). We say that \( T \in T(\mathcal{H})^+ \) is pure if it is an extreme element of this set. Equivalently, \( T \) is a pure state if and only if \( T = P[\varphi] \) for some unit vector \( \varphi \). The spectral theorem for compact operators assures that any state can be decomposed as a \( \sigma \)-convex combination of pure states.

2.2. For any state \( T \in T(\mathcal{H})^+ \) we have the following inclusions of subspaces:

\[
\text{ran}(T) \subseteq \text{ran}(T^{1/2}) \subseteq \overline{\text{ran}}(T^{1/2}) = \overline{\text{ran}}(T).
\]  

(2)

It is a restatement of a classical result for compact operators that these inclusions are equalities if and only if \( T \) is of finite rank.

2.3. A vector state \( \varphi \) is a (convex) component of a state \( T \) whenever \( T = \lambda P[\varphi] + (1 - \lambda)T' \) for some \( 0 < \lambda \leq 1 \) and for some \( T' \in T(\mathcal{H})^+ \). The state \( P[\varphi] \) is a component of \( T \) if and only if \( \varphi \in \text{ran}(T^{1/2}) \) [4]. The physical interpretation of the subspace \( \text{ran}(T^{1/2}) \) rests on this result. A decomposition \( T = \sum \lambda_i P[\varphi_i] \) of a mixed state \( T \) into vector states \( P[\varphi_i] \) is irreducible when for each \( i, \varphi_i \notin \text{lin}\{\varphi_1, \ldots, \varphi_{i-1}, \varphi_{i+1}, \ldots\} \). We say that a vector state \( P[\varphi] \) is an irreducible component of \( T \) if \( P[\varphi] \) participates an irreducible decomposition of \( T \). According to Hadjisawas [4], \( P[\varphi] \) is an irreducible component of \( T \) if and only if \( \varphi \in \text{ran}(T) \). Again, this is the property which gives the physical interpretation of the subspace \( \text{ran}(T) \).

2.4. Let \( P_T \) denote the support projection of \( T \), that is,

\[
P_T = P[\ker(T)^+] = P[\overline{\text{ran}}(T)].
\]  

(3)

We recall that \( P_T \) is the smallest projection operator such that \( T = TP_T = P_T T \). Let \( \mathcal{P}(\mathcal{H}) \) denote the set of projection operators on \( \mathcal{H} \). For any state \( T \) we define

\[
P_1(T) := \{P \in \mathcal{P}(\mathcal{H}) : \text{Tr}[TP] = 1\} = \{P \in \mathcal{P}(\mathcal{H}) : P_T \leq P\},
\]  

(4)

\[
P_1(T) := \{P \in \mathcal{P}(\mathcal{H}) : P \land P_T \neq 0\} = \cup\{P_1(P[\varphi]) : \varphi \in \overline{\text{ran}}(T)\}.
\]  

(5)

Clearly \( P_1(T) \subseteq \mathcal{P}_1(T) \) for all \( T \in T(\mathcal{H})^+ \), and \( P_1(T) = \mathcal{P}_1(T) \) if and only if \( T \) is a pure state. Moreover, if \( P \in \mathcal{P}_1(T) \), then \( \text{Tr}[TP] \neq 0 \), but not the converse (as shown explicitly by an example in Subsection 3.3). This result is to be confronted with the fact that for pure states \( T, P \in \mathcal{P}_1(T) \) if and only if \( \text{Tr}[TP] = 1 \).

3. Measurements

3.1. When modelling a measurement of an observable, we can restrict ourselves to consider a discrete observable \( A = \sum a_i P_i \). This is not an essential restriction, since we can approximate any observable by a discrete one, so that the probability distributions (1) for the measurement outcomes of the two observables are as close as we like. To model a measurement of \( A \), one fixes a measuring apparatus, with its Hilbert space \( \mathcal{K} \), an initial state \( \Phi \) of the apparatus, a pointer observable \( Z = \sum z_i Z_i \), and a (unitary) measurement coupling \( U : \mathcal{H} \otimes \mathcal{K} \rightarrow \mathcal{H} \otimes \mathcal{K} \). If \( P[\varphi] \) is the initial state of the measured system, then \( P[U(\varphi \otimes \Phi)] \) is the system–apparatus state after
the measurement. Denoting the corresponding reduced states of the measured system and the measuring apparatus as $T(\varphi)$ and $W(\varphi)$, respectively, we have the schematic representation of the state transformations associated with a measurement:

$$
\begin{array}{c}
P[\varphi] \\
\downarrow \\
P[\varphi \otimes \Phi] \rightarrow P[U(\varphi \otimes \Phi)] \\
\downarrow \\
P[\Phi] \\
\end{array}
$$

$$
\begin{array}{c}
\rightarrow T(\varphi) \\
\rightarrow W(\varphi)
\end{array}
$$

A minimal requirement for $K, Z, \Phi,$ and $U$ to be considered as a mathematical schematization of a measurement of $A$ is given by the three equivalent conditions [2]:

1. for any $i$ and $\varphi$, if $p_{a_i}(a_i) = 1$, then $p_{W(\varphi)}(z_i) = 1$; (6a)
2. for any $i$ and $\varphi$, if $P_i \in P_i(P[\varphi])$, then $Z_i \in P_i(W(\varphi))$; (6b)
3. for any $i$ and $\varphi$, $p^2_{\Phi}(a_i) = p^2_{W(\varphi)}(z_i)$. (6c)

3.2. To determine the structure of a measurement $\langle K, Z, \Phi, U \rangle$ of an observable $A = \sum a_i P_i$, we assume that the pointer observable is minimal in the sense that it is just sufficiently big to distinguish among different values $a_i$ of $A$. Any choice $Z = \sum z_i Z_i \equiv \sum z_i P[\Phi_i]$ will then do, where $\{\Phi_i\}$ is a basis of $K$. With that choice all unitary mappings $U$ which make $\langle K, Z, \Phi, U \rangle$ satisfy conditions (6) are completely characterized [2]. To describe these solutions, we assume here, for notational simplicity only, that all the eigenvalues of $A$ are nondegenerate: $A = \sum a_i P_i = \sum a_i P[\varphi_i]$, where $\{\varphi_i\}$ is a basis of eigenvectors of $A$, $A \varphi_i = a_i \varphi_i$. If $\langle K, Z, \Phi, U \rangle$ is a measurement of $A$, then

$$
\text{for each } \varphi_i, \quad U(\varphi_i \otimes \Phi) = \gamma_i \otimes \Phi_i,
$$

where $\gamma_i$ are unit vectors $\gamma_i = \sum_j \langle \varphi_j \otimes \Phi_i | U(\varphi_i \otimes \Phi) | \varphi_j \rangle \varphi_j$. On the other hand, given any set of unit vectors $\{\gamma_i\}$ of $\mathcal{H}$, then (7) extends to a unitary mapping $U$ such that $\langle K, Z, \Phi, U \rangle$ is a measurement of $A$. The final states $T(\varphi)$ of the measured system and $W(\varphi)$ of the apparatus are then given as

$$
\begin{align*}
T(\varphi) &= \sum_i |\langle \varphi | \varphi_i \rangle|^2 P[\gamma_i], \\
W(\varphi) &= \sum_{ij} \langle \varphi_i | \varphi \rangle \langle \varphi | \varphi_j \rangle \langle \gamma_j | \gamma_i \rangle |\Phi_i \rangle |\Phi_j\rangle.
\end{align*}
$$

We stress that conditions (6) pose no restrictions on the measurement coupling $U$ generating unit vectors $\{\gamma_i\}$.

The best known choice of the set $\{\gamma_i\}$ is, of course, $\{\varphi_i\}$, in which case (7) takes the canonical form of a von Neumann–Lüders measurement:

$$
U_{NL}(\varphi_i \otimes \Phi) = \varphi_i \otimes \Phi_i.
$$

3.3. Conditions (6) do not exhaust the physics of measurement process. Other requirements are called for. In particular, there are the following requirements which rely
PROPERTIES OF THE RANGE OF A STATE OPERATOR

on the structure of the range of $W(\varphi)$:

for any $i$ and $\varphi$, if $p_\varphi^A(a_i) \neq 0$, then $\Phi_i \in \overline{\text{ran}}(W(\varphi))$; \hfill (10)

for any $i$ and $\varphi$, if $p_\varphi^A(a_i) \neq 0$, then $\Phi_i \in \text{ran}(W(\varphi)^{1/2})$; \hfill (11)

for any $i$ and $\varphi$, if $p_\varphi^A(a_i) \neq 0$, then $\Phi_i \in \text{ran}(W(\varphi))$. \hfill (12)

In view of the subspace inclusions (2), the mathematical relevance of these requirements is obvious. In addition, each of them has an important physical content. Condition (10) derives its physical meaning from the following equivalent condition:

for any $i$ and $\varphi$, if $p_\varphi^A(a_i) \neq 0$, then $P[\Phi_i] \in P_1(W(\varphi))$,

which is at the heart of the Copenhagen variant of the modal interpretation of quantum mechanics [3, 5]. Conditions (11) and (12) obtain their physical content from the properties of decomposability of a mixed state discussed in Subsection 2.3. Indeed, a pointer state $P[\Phi_i]$ is a component of the final apparatus state $W(\varphi)$ exactly when $\Phi_i$ is in the range of the square root of $W(\varphi)$. On its turn, condition $\Phi_i \in \text{ran}(W(\varphi))$ is necessary and sufficient for $P[\Phi_i]$ to be an irreducible component of $W(\varphi)$.

Before turning to study the properties (10)–(12), we give an example which shows that the above requirements go, indeed, beyond the minimal conditions (6). Let $\gamma_k = \gamma_l$ for some $k$ and $l$, and consider an initial state $P[\varphi] = P[\frac{1}{\sqrt{2}}(\Phi_k \pm \Phi_l)]$. The state of the apparatus after the measurement is $W(\varphi) = P[\frac{1}{\sqrt{2}}(\Phi_k + \Phi_l)]$, so that its support projection is the state (one-dimensional projection) itself, and neither $\Phi_k$ nor $\Phi_l$ is contained in $\text{ran}(W(\varphi))$. Equivalently, neither $P[\Phi_k]$ nor $P[\Phi_l]$ is in $P_1(W(\varphi))$, though both of the measurement outcome probabilities $p_\varphi^A(a_k)$ and $p_\varphi^A(a_l)$ are now nonzero.

4. The closure of the range of $W(\varphi)$

Consider the measurement of $A$ with a minimal pointer observable $Z = \sum z_i P[\Phi_i]$ and let $(\gamma_i) \subset \mathcal{H}$ be the measurement coupling generating unit vectors according to (7). The final state $W(\varphi)$ of the apparatus is then given by (9). Since we are interested in characterizing the closure of the range of this state for all possible initial states $\varphi$ of the measured object, we may assume that $\langle \varphi \mid \varphi_i \rangle \neq 0$ for each $i$. For such states the following conditions are equivalent:

\begin{align}
\Phi_i \in \overline{\text{ran}}(W(\varphi)) & \quad \text{for all } i; \quad \text{(13a)} \\
\text{ran}(W(\varphi)) = K & \quad \text{; (13b)} \\
\ker(W(\varphi)) = \{0\} & \quad \text{. (13c)}
\end{align}

We start with characterizing the kernel of the state $W(\varphi)$.

4.1. From (9) we get that for each $\Psi \in K$

\begin{equation}
W(\varphi)\Psi = \sum_{i,j} \langle \varphi_i \mid \varphi \rangle \langle \varphi \mid \varphi_j \rangle \langle \gamma_j \mid \gamma_i \rangle \langle \Phi_j \mid \Psi \rangle \Phi_i.
\end{equation}
Therefore,

\[ W(\varphi)\Psi = 0 \Rightarrow \sum_{j=1}^{\infty} \langle \varphi | \varphi_j \rangle \langle \gamma_j | \gamma_i \rangle \langle \Phi_j | \Psi \rangle = 0 \quad \forall i \]

\[ \Rightarrow \langle (\sum_{j=1}^{\infty} \langle \varphi_j | \varphi \rangle \langle \Phi_j | \gamma_j \rangle) | \gamma_i \rangle = 0 \quad \forall i \]

\[ \Rightarrow \sum_{j=1}^{\infty} \langle \varphi_j | \varphi \rangle \langle \Psi | \Phi_j \rangle \gamma_j \perp \overline{\text{lin}(\gamma_i)}. \]

On the other hand,

\[ \sum_{j=1}^{\infty} \langle \varphi_j | \varphi \rangle \langle \Psi | \Phi_j \rangle \gamma_j \in \overline{\text{lin}(\gamma_i)}, \]

so that

\[ \sum_{j=1}^{\infty} \langle \varphi_j | \varphi \rangle \langle \Psi | \Phi_j \rangle \gamma_j = 0. \]

Conversely, if \( \Psi \in \mathcal{K} \) is such that

\[ \sum_{j=1}^{\infty} \langle \varphi_j | \varphi \rangle \langle \Psi | \Phi_j \rangle \gamma_j = 0, \]

then a direct calculation shows that \( \Psi \in \ker(W(\varphi)) \). In conclusion:

\[ \ker(W(\varphi)) = \{ \Psi \in \mathcal{K} : \sum_{j=1}^{\infty} \langle \varphi_j | \varphi \rangle \langle \Psi | \Phi_j \rangle \gamma_j = 0 \} \]

(with the assumption that \( \langle \varphi_j | \varphi \rangle \neq 0 \) for all \( j \)).

4.2. The example at the end of Subsection 3.3 suggests that the linear independence of the vectors \( \gamma_i \) is necessary for \( \Phi \) to be in \( \text{ran}(W(\varphi)) \). Indeed, it turns out that this is the case. Moreover, one can show, by a direct computation, that in the finite dimensional case this is also a sufficient condition. However, the following example shows that, in general, the linear independence of the vectors \( \gamma_i \) is not sufficient for (13). To witness, define

\[ \gamma_1 = \sum_{i=1}^{\infty} \frac{1}{i^2} \varphi_i, \quad \gamma_{n+1} = \varphi_n, \quad n = 1, 2, \ldots \]

Apart from the normalization (\( \| \gamma_1 \| = \pi^2/6 \)), these vectors define an \( A \)-measurement (since any set does). Clearly, the set \( (\gamma_i) \) is linearly independent. Let \( \varphi = \sum_{n=1}^{\infty} (1/n) \varphi_n \), so that \( \langle \varphi | \varphi_n \rangle = 1/n \neq 0 \) for all \( n \). Define a sequence \((a_k)\) such that \( a_1 = -1 \), and \( a_k = 1/(k-1)^2 \) for \( k > 1 \). Then the series \( \sum_{k=1}^{\infty} k^2 |a_k|^2 \) is convergent. Hence \( \Psi = \sum_{k=1}^{\infty} a_k k \Phi_k \in \mathcal{K} \), and \( \Psi \neq 0 \). But now
\[
\sum_{j=1}^{\infty} \langle \varphi_j | \varphi \rangle \langle \Psi | \Phi_j \rangle \gamma_j = -\sum_{n=1}^{\infty} \frac{1}{n^2} \varphi_n + \sum_{j=2}^{\infty} \frac{1}{j(j-1)^2} \varphi_{j-1} \\
= -\sum_{n=1}^{\infty} \frac{1}{n^2} \varphi_n + \sum_{j=1}^{\infty} \frac{1}{j^2} \varphi_j = 0,
\]
which shows that \( \Psi \in \ker(W(\varphi)) \), that is, \( \text{ran}(W(\varphi)) \neq \mathcal{K} \).

4.3. The above example shows that the algebraic condition of linear independence of the vectors \( (\gamma_i) \) is to be strengthened into something which takes care of "infinite linear combinations", like \( \gamma_1 = \sum_i 1/i^2 \gamma_{i+1} \), above. Therefore we adopt the following definition:

A set of vectors \( (\gamma_i) \in \mathcal{H} \) is \( \ell_1 \)-linearly independent if for each \( \ell \)-sequence \( (a_i) \) the condition \( \sum_{i=1}^{\infty} a_i \gamma_i = 0 \) implies that \( a_i = 0 \) for all \( i = 1, 2, \ldots \).

We recall that if \( (a_i) \in \ell_1 \), then the series \( \sum_{i=1}^{\infty} a_i \gamma_i \) is convergent in \( \mathcal{H} \), independently of the choice of the set \( (\gamma_i) \).

We come now to the main result of this section.

4.4. THEOREM. \( \text{ran}(W(\varphi)) = \mathcal{K} \) if and only if the set \( (\gamma_i) \) is \( \ell_1 \)-linearly independent.

**Proof:** Assume first that \( (\gamma_i) \) is \( \ell_1 \)-linearly independent. Let \( \Psi \in \ker(W(\varphi)) \), so that \( \sum_{j=1}^{\infty} \langle \varphi_j | \varphi \rangle \langle \Psi | \Phi_j \rangle \gamma_j = 0 \). Since \( \langle \varphi_j | \varphi \rangle \) and \( \langle \Psi | \Phi_j \rangle \) are \( \ell_2 \)-sequences, \( \langle \varphi_j | \varphi \rangle \langle \Psi | \Phi_j \rangle \) is an \( \ell_1 \)-sequence. Therefore, \( \langle \varphi_j | \varphi \rangle \langle \Psi | \Phi_j \rangle = 0 \) for all \( j \). Since \( \langle \varphi_j | \varphi \rangle \neq 0 \), this means that \( \langle \Psi | \Phi_j \rangle = 0 \) for all \( j \), that is, \( \Psi = 0 \). Hence, \( \ker(W(\varphi)) = \{0\} \), and thus \( \text{ran}(W(\varphi)) = \mathcal{K} \).

Conversely, suppose that \( \text{ran}(W(\varphi)) = \mathcal{K} \). Let \( (a_i) \in \ell_1 \) be such that \( \sum_{i=1}^{\infty} a_i \gamma_i = 0 \). Define

\[
b_n = |a_n|^{1/2} e^{\frac{i}{2} \arg a_n}, \quad n = 1, 2, \ldots
\]

Then \( (b_n) \in \ell_2 \), and \( b_n^2 = a_n \) for all \( n \). Define also

\[
c_n = b_n \quad \text{if} \quad b_n \neq 0,
\]
\[
c_n = \frac{1}{n} \quad \text{if} \quad b_n = 0.
\]

Therefore, \( (c_n) \in \ell_2 \) and \( c_n \neq 0 \) for all \( n \). Let \( \varphi = \sum_i c_i \varphi_i \in \mathcal{H} \). Hence \( \langle \varphi | \varphi_i \rangle = c_i \neq 0 \) for each \( i \). Finally, let \( \Psi = \sum_{i=1}^{\infty} b_i \Phi_i \in \mathcal{K} \). With this choice we get

\[
\sum_{i=1}^{\infty} \langle \varphi_i | \varphi \rangle \langle \Psi | \Phi_i \rangle \gamma_i = \sum_{i=1}^{\infty} c_i b_i \gamma_i = \sum_{i=1}^{\infty} a_i \gamma_i = 0,
\]
which now implies that \( \Psi \in \ker(W(\varphi)) \). Therefore \( \Psi = 0 \). But then \( b_i = 0 \) for all \( i \), and thus also \( a_i = 0 \) for all \( i \), which is to say that \( (\gamma_i) \) is \( \ell_1 \)-linearly independent.

4.5. The above theorem gives necessary and sufficient conditions for the measurement of \( A \), with a minimal pointer observable \( Z = \sum_i z_i P[\Phi_i] \), to be such that whenever the measurement outcome probability \( p_{\varphi}^A(a_i) \) is nonzero, then also \( P[\Phi_i] \in \mathcal{P}_1(W(\varphi)) \).
These conditions are given in terms of the measurement coupling $U$ generating unit vectors $(\gamma_i)$. This result forms the mathematical basis for the Copenhagen variant of the modal interpretation of quantum mechanics [3, 5].

5. The range of the square root of $W(\varphi)$

Next we study the conditions under which pointer states $P[\Phi_i]$ are possible components of the final apparatus state $W(\varphi)$. Again, we assume that $\langle \varphi | \varphi_i \rangle \neq 0$ for all $i$.

Let $(\gamma_i)$ be a sequence of unit vectors of $\mathcal{H}$. Introducing for each $i$ the notations
\[ V_i^0 = \text{lin}\{\gamma_1, \ldots, \gamma_{i-1}, \gamma_{i+1}, \ldots\}, \]
\[ V_i = \overline{\text{lin}}(V_i^0), \]
\[ P_i \text{ the projection on } V_i, \]
we come to the main theorem of this section.

THEOREM. The following two conditions are equivalent:
(a) $\Phi_i \in \text{ran}(W(\varphi)^{1/2})$ for all $i$;
(b) $\gamma_i \notin V_i$ for all $i$.

Remark: Note that the second condition of this theorem means that the decomposition $T(\varphi) = \sum_i |\langle \varphi | \varphi_i \rangle|^2 P[\gamma_i]$ of the final object state is irreducible. Therefore, this theorem shows that pointer states $P[\Phi_i]$ are possible components of the final apparatus state $W(\varphi)$ exactly when the measurement coupling $U$ generating vectors $(\gamma_i)$ form an irreducible decomposition of the final object state $T(\varphi)$. We emphasise that this does not, however, mean that $(\gamma_i)$ are orthogonal.

For sake of clarity, the proof of the above theorem is organized in three lemmas and a few technical remarks. It is not a loss of generality to assume that $\mathcal{H} = \overline{\text{lin}}\{\gamma_1, \ldots, \gamma_i, \ldots\}$. Without this assumption the proof would remain essentially unchanged but we should have to introduce some notational complications.

5.1. We introduce the following definition: a sequence $(\theta_i)$ of vectors of $\mathcal{H}$ is a dual sequence of $(\gamma_i)$ if $\langle \theta_i | \gamma_j \rangle = \delta_{ij}$.

**LEMMA.** A sequence $(\gamma_i)$ has a dual sequence if and only if $\gamma_i \notin V_i$ for all $i$.

**Proof:** Suppose that $(\gamma_i)$ has a dual sequence $(\theta_i)$. Then $\theta_i \perp V_i$, since $\langle \theta_i | \gamma_j \rangle = 0$ if $j \neq i$. Suppose, for absurd, that $\gamma_i \in V_i$, so that $\langle \theta_i | \gamma_i \rangle = 0$. By the definition of a dual sequence, $\langle \theta_i | \gamma_i \rangle = 1$. Therefore, we have a contradiction and thus $\gamma_i \notin V_i$. Conversely, suppose that $\gamma_i \notin V_i$. This implies that $V_i^\perp = V_i^{0\perp} \neq \{0\}$, and we thus have the nontrivial decomposition: $\mathcal{H} = V_i \oplus V_i^{0\perp}$. By assumption, the vector $\gamma_i - P_i \gamma_i$ is nonzero. Since $P_i(\gamma_i - P_i \gamma_i) = P_i \gamma_i - P_i \gamma_i = 0$, $\gamma_i - P_i \gamma_i \in V_i^{0\perp}$. Moreover,
\[ \langle \gamma_i | \gamma_i - P_i \gamma_i \rangle = \langle \gamma_i | \gamma_i \rangle - \langle \gamma_i | P_i \gamma_i \rangle \neq 0. \]
We may thus define

\[ \theta_i = \frac{\gamma_i - P_1\gamma_i}{\langle \gamma_i | \gamma_i - P_2\gamma_i \rangle} \]

Then we have \( \langle \theta_i | \gamma_j \rangle = \delta_{ij} \).

**Remark:** The sequence \( (\theta_i) \) dual to \( (\gamma_i) \) is *unique*, that is, it is uniquely determined by \( (\gamma_i) \). In fact, suppose \( (\theta_i') \) is another dual sequence, then \( \langle \theta_i - \theta_i' | \gamma_j \rangle = 0 \) for all \( j \), also for \( i = j \). Since \( \text{lin}\{\gamma_1, \ldots, \gamma_i, \ldots\} = \mathcal{H} \), we then get \( \theta_i - \theta_i' = 0 \).

**Remark:** The subspace \( V_i^0 = V_i^\perp \) is *one-dimensional*, and it is generated by the vector \( \theta_i \). Indeed, let \( \psi_i \) be any vector (\( \neq 0 \)) in \( V_i^\perp \); \( \psi_i \) cannot be orthogonal to \( \gamma_i \), since it would imply that \( \langle \gamma_j | \psi_i \rangle = 0 \) for all \( j \), that is, \( \psi_i \equiv 0 \). Hence we define \( \theta_i' = \psi_i / \langle \gamma_i | \psi_i \rangle \). Consider then \( \theta_i - \theta_i' \); we have \( \langle \theta_i - \theta_i' | \gamma_j \rangle = 0 \) for all \( j \), also for \( i = j \), so that \( \theta_i = \theta_i' \).

**Remark:** The previous point shows that any \( \psi \in \mathcal{H} \) has the decomposition \( \psi = a\theta_i + P_1\psi_i \), where \( a = \langle \theta_i | \psi_i \rangle / \| \theta_i \|^2 \).

5.2. We denote

\[ \Omega_i := \frac{1}{\langle \varphi | \varphi_i \rangle} W(\varphi)^{1/2}\Phi_i. \]

Then, by (9),

\[ \langle \Omega_i | \Omega_j \rangle = \frac{1}{\langle \varphi_i | \varphi_j \rangle \langle \varphi | \varphi_j \rangle} \langle \Phi_i | W(\varphi)\Phi_j \rangle = \langle \gamma_j | \gamma_i \rangle. \]

**Lemma.** There exists a sequence \( (\Omega_i) \subset \mathcal{K} \) such that \( W(\varphi)^{1/2}\Psi_i = \Phi_i \) for all \( i \) if and only if \( (\Omega_i) \) has a dual sequence.

**Proof:** Suppose that there is a sequence \( (\Psi_i) \) such that \( W(\varphi)^{1/2}\Psi_i = \Phi_i \) for all \( i \). Then

\[ \langle \langle \varphi_i | \varphi \rangle \Psi_i | \Omega_j \rangle = \frac{\langle \varphi | \varphi_i \rangle}{\langle \varphi_i | \varphi_j \rangle} \langle \Psi_i | W(\varphi)^{1/2}\Phi_j \rangle = \delta_{ij}. \]

Hence \( (\Omega_i) \) has a dual sequence. Conversely, suppose that \( (\Omega_i) \) has a dual sequence \( (\Gamma_i) \) and define \( \Psi_i = \Gamma_i / \langle \varphi_i | \varphi \rangle \). Then

\[ \langle W(\varphi)^{1/2}\Psi_i | \Phi_j \rangle = \langle \Psi_i | W(\varphi)^{1/2}\Phi_j \rangle = \frac{\langle \varphi | \varphi_i \rangle}{\langle \varphi_i | \varphi_i \rangle} \langle \Gamma_i | \Omega_j \rangle = \delta_{ij}. \]

Since \( (\Phi_j) \) is an orthonormal basis of \( \mathcal{K} \), this shows that \( W(\varphi)^{1/2}\Psi_i = \Phi_i \).

5.3. **Lemma.** The sequence \( (\Omega_i) \) has a dual sequence (denoted \( (\Gamma_i) \)) if and only if \( (\gamma_i) \) has a dual sequence (denoted \( (\theta_i) \)).
Proof: Suppose that \((\Omega_i)\) has a dual sequence. Then by the previous lemma, \(\Phi_i \in \operatorname{ran}(W(\varphi)^{1/2})\) for all \(i\), and \(W(\varphi)^{1/2}\) is injective.

The sequence of vectors \((\Omega_i)\) is linearly independent. In fact, if we have

\[
0 = \sum_{i=1}^{n} a_i \Omega_i = W(\varphi)^{1/2} \left( \sum_{i=1}^{n} \frac{a_i}{\varphi_i} \Phi_i \right),
\]

the injectivity of \(W(\varphi)^{1/2}\) implies that \(\sum_{i=1}^{n} (a_i/\langle \varphi | \varphi_i \rangle) \Phi_i = 0\), and therefore \(a_i = 0\) for all \(i\). We define the mapping \(U: \operatorname{lin}\{\Omega_1, \ldots, \Omega_i, \ldots\} \to \mathcal{H}\) by

\[
U\Psi = U \left( \sum_{i=1}^{M} a_i \Omega_i \right) := \sum_{i=1}^{M} a_i \gamma_i.
\]

\(U\) is well defined since \((\Omega_i)\) is an algebraic basis of \(\operatorname{lin}\{\Omega_1, \ldots, \Omega_i, \ldots\}\); \(U\) is an antilinear isometric mapping, in fact

\[
\|U\Psi\|^2 = \sum_{i,j} a_i \overline{a_j} \langle \gamma_i | \gamma_j \rangle = \sum_{i,j=1}^{M} a_i \overline{a_j} \langle \Omega_j | \Omega_i \rangle = \|\Psi\|^2.
\]

\(U\) can be uniquely extended to an antilinear isometric mapping \(\overline{\operatorname{lin}\{\Omega_1, \ldots, \Omega_i, \ldots\}} \to \mathcal{H}\), which we still denote by \(U\). Notice that \(\overline{\operatorname{lin}\{\Omega_1, \ldots, \Omega_i, \ldots\}} = \overline{\operatorname{ran}(W(\varphi)^{1/2})} = \mathcal{K}\).

Define \(\theta_i = U\Gamma_i\), and recall that, by definition, \(\gamma_i = U\Omega_i\). Therefore

\[
\langle \theta_i | \gamma_j \rangle = \langle U\Gamma_i | U\Omega_j \rangle = \langle \Omega_j | \Gamma_i \rangle = \delta_{ji},
\]

since \(U\) is isometric. This shows that \((\gamma_i)\) has a dual sequence. The proof of the reverse implication is exactly analogous, and we omit it here.

6. The range of \(W(\varphi)\)

Our final result characterizes \(A\)-measurements for which the pointer states \(\Phi_i\) are possible irreducible convex components of the final apparatus state \(W(\varphi)\). This is the subject of the next theorem.

**Theorem.** The following two conditions are equivalent:

(a) \(\Phi_i \in \operatorname{ran}(W(\varphi))\) for all \(i\);

(b) the set \((\gamma_i)\) is linearly independent and \(P_i \gamma_i \in V_i^0\) for all \(i\).

Before entering the proof of this theorem, we note that the condition \(P_i \gamma_i \in V_i^0\) means that \(\gamma_i = a\theta_i + \sum_{j=1, j \neq i}^{M} a_j \gamma_j, \theta_i \in V_i^\perp\).

**Proof:** Assume first that \(\Phi_i \in \operatorname{ran}(W(\varphi))\) for all \(i\), and let \((\Psi_i) \subset \mathcal{K}\) be a sequence of vectors for which \(W(\varphi)\Psi_i = \Phi_i\). Since \((\langle \varphi_j | \varphi \rangle)\) and \((\langle \Psi_i | \Phi_j \rangle)\) are \(\ell_2\)-sequences for all \(i\), then \((\langle \varphi_j | \varphi \rangle \langle \Psi_i | \Phi_j \rangle)\) are \(\ell_1\) sequences for all \(i\). This shows that the series

\[
\sum_{j=1}^{\infty} \langle \varphi_j | \varphi \rangle \langle \Psi_i | \Phi_j \rangle \gamma_j
\]
is convergent in $\mathcal{H}$. A quick calculation shows that the sequence $(\theta_i)$, where

$$\theta_i = \langle \varphi | \varphi_i \rangle \sum_j \langle \varphi_j | \varphi \rangle \langle \Psi_i | \Phi_j \rangle \gamma_j,$$

is dual of the sequence $(\gamma_i)$ and, hence, the set $(\gamma_i)$ is linearly independent.

Since $\langle \gamma_j | \theta_k \rangle = \delta_{jk}$, we have further that

$$\frac{\langle \theta_k | \theta_k \rangle}{\langle \varphi_i | \varphi \rangle \langle \varphi | \varphi_k \rangle} = \langle \Phi_k | \Psi_i \rangle$$

and, hence, that

$$\gamma_i = \frac{1}{\langle \theta_i | \theta_i \rangle^2} \theta_i - \frac{1}{\langle \theta_i | \theta_i \rangle^2} \sum_{j \neq i} \langle \theta_j | \theta_i \rangle \gamma_j.$$

Therefore, the only thing to be shown is that the last sum contains only a finite number of terms $\neq 0$. For this, observe the following facts: the sequence $((\theta_i | \theta_k))_k$ is bounded by $\|\Psi_i\|$ (it follows by Cauchy Schwarz inequality applied to (14)) and independent of $\varphi$ (since $(\theta_i)$ is the dual sequence of $(\gamma_i)$); since the sequence $((\Phi_k | \Psi_i))_k$ is in $\ell_2$ for each fixed $i$, the sequence

$$\left( \frac{\langle \theta_i | \theta_k \rangle}{\langle \varphi_i | \varphi \rangle \langle \varphi | \varphi_k \rangle} \right)_k$$

is also an $\ell_2$-sequence for each $i$ and $\varphi$. Fixing $i$, we define $\overline{\varphi}$ such that

$$\langle \varphi_k | \overline{\varphi} \rangle = \frac{\langle \theta_k | \theta_i \rangle}{k} \quad \text{when } \langle \theta_i | \theta_k \rangle \neq 0,$$

$$\langle \varphi_k | \overline{\varphi} \rangle = \frac{1}{k}, \quad \text{otherwise.}$$

Since $((\theta_i | \theta_k))_k$ is bounded, $\overline{\varphi}$ is well-defined and $\neq 0$. We choose $\varphi$ as $\overline{\varphi}/\|\overline{\varphi}\|$ so that we have

$$\frac{\langle \theta_i | \theta_k \rangle}{\langle \varphi_i | \varphi \rangle \langle \varphi | \varphi_k \rangle} = \frac{k}{\langle \varphi_i | \varphi \rangle \langle \varphi | \varphi_k \rangle} \|\overline{\varphi}\|$$

for each $k$ such that $\langle \theta_i | \theta_k \rangle \neq 0$ ($i$ is fixed); then $\langle \theta_i | \theta_k \rangle \neq 0$ only for a finite number of $k$ (for each fixed $i$).

To prove the converse implication, we observe that $\gamma_i \notin V_i$ because $\gamma_i$ are linearly independent and $P_i \gamma_i \in V_i^0$; hence, due to the Lemma in Subsection 5.1, we have that $(\gamma_i)$ has a dual sequence $(\theta_i)$. Moreover, the assumption $P_i \gamma_i \in V_i^0$ now shows that

$$\theta_i = \sum_{j=1}^M a_k \gamma_k = \sum_{k=1}^M (\theta_k, \theta_i) \gamma_k,$$

where the sum is finite. From this relation we have that for each $i$ only a finite number of products $\langle \theta_i | \theta_k \rangle$ is $\neq 0$. For each fixed $i$ define $\Psi_i \in \mathcal{K}$ such that

$$\langle \Phi_k | \Psi_i \rangle = \frac{\langle \theta_i | \theta_k \rangle}{\langle \varphi_i | \varphi \rangle \langle \varphi | \varphi_k \rangle}.$$
$\Psi_k$ is well defined and a direct computation shows that $W(\varphi)\Psi_i = \Phi_i$. This ends the proof.

In addition to the above requirements one could still ask whether all the pointer states $\Phi_i$, for which $p_{\varphi}(a_i) \neq 0$, may occur simultaneously as the components of $W(\varphi)$, that is, $W(\varphi) = \sum p_{\varphi}(a_i) P[\Phi_i]$. This is the case exactly when the vectors $\gamma_i$ are mutually orthogonal. The decompositions $T(\varphi) = \sum p_{\varphi}^2(a_i) P[\gamma_i]$ and $W(\varphi) = \sum p_{\varphi}^2(a_i) P[\Phi_i]$ both are then the spectral ones. Various characterizations of measurements leading to this situation are known [1, 2]. Typically, the value reproducibility and the strong correlation properties of measurements are such.

7. Properties of bases in Banach spaces

The main theme of this paper has been to express properties of the range of a state operator in terms of properties of the sequence $(\gamma_i)$. Each of those properties which were considered in the previous sections can be viewed as an "infinite generalization" of the algebraic notion of linear independence of the vectors $(\gamma_i)$. It turns out that these properties arise in a natural way in the theory of bases in Banach spaces, that has been a guiding line for our study. In the sequel we shall very briefly explain the main ideas of this theory [6].

Let $(\gamma_i) \subset \mathcal{H}$ be a sequence of unit vectors such that $\mathcal{H} = \overline{\text{lin}}\{\gamma_i\}$. For the moment we consider $\mathcal{H}$ only as a Banach space. One says that $(\gamma_i)$ is a Schauder basis of $\mathcal{H}$ if for any $\varphi \in \mathcal{H}$ there exists a unique sequence $(a_i)$ of complex numbers such that $\varphi = \sum a_i \gamma_i$. (Of course, if $(\gamma_i)$ is an orthogonal sequence, it is a Hilbert basis of $\mathcal{H}$, which is a very special kind of Schauder basis for Hilbert spaces).

If $(\gamma_i)$ is a Schauder basis, we can define the "coordinate forms" $f_i$ by $f_i(\varphi) := a_i$, if $\varphi = \sum a_i \gamma_i$. The mappings $f_i$ are linear forms on $\mathcal{H}$ and it is a classical result that they are continuous on $\mathcal{H}$. The continuity of these forms has important equivalent characterizations. To explain them, we recall the following notion from Section 5: A sequence $(\omega_i)$ of elements of the (topological) dual of $\mathcal{H}$ is a dual sequence of $(\gamma_i)$ if $\omega_i(\gamma_j) = \delta_{ij}$. (In Section 5 this notion was given in the Hilbert space context). It is a standard result in Banach spaces that the following conditions are equivalent:

(a) $f_i$ is continuous for all $i$;
(b) $(\gamma_i)$ has a dual sequence;
(c) $\gamma_i \notin V_i$ for all $i$.

If we use the Hilbert space properties of $\mathcal{H}$, it is an elementary exercise to show that $f_i$ are continuous if and only if $(\gamma_i)$ has a dual sequence. In Section 5 we have given a direct proof of the fact that $(\gamma_i)$ has a dual sequence if and only if $\gamma_i \notin V_i$ for all $i$. We stress that these equivalences can also be proved in the general setting of Banach spaces [6].

The continuity of coordinate forms $f_i$ implies that the sequence $(\gamma_i)$ is topologically linearly independent. This means that

$$\sum a_i \gamma_i = 0 \iff a_i = 0 \text{ for all } i.$$
In turn, topological linear independence of the sequence $(\gamma_i)$ implies its $\ell_1$-linear independence (cf. Subsection 4.3).

In conclusion, we have the following situation:

\begin{align*}
(\gamma_i) \text{ is an orthonormal basis} & \Rightarrow (15) \\
(\gamma_i) \text{ is a Schauder basis} & \Rightarrow (16) \\
(\gamma_i) \text{ has a dual sequence} & \Rightarrow (17) \\
(\gamma_i) \text{ is topologically linearly independent} & \Rightarrow (18) \\
(\gamma_i) \text{ is } \ell_1\text{-linearly independent.} & (19)
\end{align*}

None of these implications can be reversed.

Properties (16)–(18) are widely considered in the context of the theory of basis in Banach spaces. As an example, it is a classical question to find conditions to be added to (17) and (18) to get (16).

Motivated by the measurement theory we have considered for the set $(\gamma_i)$ the "finiteness" property (b) of the Theorem in Section 6. This is clearly a weakening of the orthonormality condition (15), and it implies (17). Nevertheless, it cannot be placed in a natural way into the theory of linearly independent infinite sets of vectors in Banach spaces.

8. Conclusions

Quantum mechanics allows one to predict the outcomes of measurements, but it excludes any obvious attempt to explain the occurrence of a particular result, the fact that a measurement leads to a result. These well-known difficulties, which are rooted in the irreducible probabilistic structure of quantum mechanics, are closely related to the structure, the interpretation and the decomposability properties, of the final apparatus state. The fact that the pointer observable $Z = \sum z_i P[\Phi_i]$ could have a value $z_i$ after the measurement, whenever $p^A(\alpha_i) \neq 0$, is expressed in the Copenhagen variant of the modal interpretation as $P[\Phi_i] \in P_1(W(\varphi))$, a condition which is now equivalent to $\Phi_i \in \text{ran}(W(\varphi))$. Theorem 4.2 gives necessary and sufficient conditions for this to be the case. These conditions are given in terms of the vectors $(\gamma_i)$ which define the measurement coupling.

It is equally natural to expect that a pointer state $\Phi_i$ is a possible component, or even a possible irreducible component of $W(\varphi)$, whenever $p^A(\alpha_i) \neq 0$. These conditions are, however, again additional constraints on the measurement coupling going beyond the above quoted requirement of the modal interpretation. The theorems of Sections 5 and 6 give necessary and sufficient conditions, respectively, for $\Phi_i \in \text{ran}(W(\varphi))$ and $\Phi_i \in \text{ran}(W(\varphi))$ to occur whenever $p^A(\alpha_i) \neq 0$. In particular, we recall that pointer states $\Phi_i$ are possible components of $W(\varphi)$, whenever the decomposition $T(\varphi) = \sum p^A(\alpha_i) P[\gamma_i]$ of the final state of the measured system is irreducible.
Acknowledgements

Part of this work was carried out during the visit of two of us (CG & PL) at Interdisciplinary Laboratory for Humanistic and Natural Sciences in Trieste, October 1993. We gratefully acknowledge the hospitality and kindness addressed to us.

REFERENCES


