COMMUTATIVE POVMS AND FUZZY OBSERVABLES

S. TWAREQUE ALI, CLAUDIO CARMELI, TEIKO HEINOSAARI, AND ALESSANDRO TOIGO

ABSTRACT. In this paper we review some properties of fuzzy observables, mainly as realized by commutative positive operator valued measures. In this context we discuss two representation theorems for commutative positive operator valued measures in terms of projection valued measures and describe, in some detail, the general notion of fuzzification. We also make some related observations on joint measurements.

1. INTRODUCTION

Fuzzy observables arise in several ways in quantum formalism. Their history goes back to the 1970s, and they were one of the first concrete examples of genuine positive operator valued measures (POVMs) in quantum mechanics. For commutative POVMs, this concept gives useful representation theorems in terms of projection valued measures (PVMs). On the other hand, fuzzification is a natural way to model reduction in the statistical description of the system. Fuzzy observables are also intimately related to the topic of joint measurements.

In this article we discuss all these aspects of fuzzy observables. In Section 2 we give a brief historical overview of fuzzy observables. Section 3 contains some basic definitions and fixes the notation. In Section 4 we present two representation theorems for commutative POVMs. In Section 5 we explain the connection of these results to the concept of fuzzy observables. Finally, in Section 6 we present some remarks related to joint measurements.

2. A QUICK HISTORICAL OVERVIEW

The appearance of positive operator valued measures in studying of the foundations of quantum mechanics followed along several lines. The axiomatization of quantum mechanics, built upon an analysis of the measurement process, as undertaken by Günther Ludwig and the Marburg school in the 1960's and 70's [Lud64, Lud70], led to the notion of generalized observables, which went beyond the orthodox concept of an observable in quantum mechanics. An observable in this more general setting turned out to be a positive operator valued measure on a Hilbert space. The σ-algebra of this measure was generally built on the set of possible values of the observable or observables, the implication being that it was theoretically possible to discuss joint measurements of incompatible observables. Thus, in orthodox quantum mechanics the observables of position and momentum are each characterized by projection valued measures, i.e., the spectral projectors of the corresponding self adjoint operators, defined on the space

of possible observed values of position and momentum, and they are not simultaneously measurable. In the more general setting, it is entirely possible to discuss the joint measurement of these two observables, using a positive operator valued measure defined on the joint set of values of classical position and momentum. Of course, in such a joint measurement neither position nor momentum can assume sharp values, implying necessarily a certain *fuzzification*. This line of thought leads to a one type of formulation of Heisenberg's uncertainty principle - a condition ensuring that mutually exclusive measurements of position and momentum can be reconciled if an appropriate trade-off is accepted. For a recent exposition of this topic, we refer to [BHL07].

The second line of development, which led to the use of POVMs was the generalization of the notion of a conditional expectation in probability theory. An algebraic formulation of the notion of conditional expectation, in the context of classical, i.e., commutative, probability theory was worked out in the 1950's [Moy54, NT54], which was further developed and extended to non-commutative probability theories, in which the classical algebra of stochastic variables was replaced by a non-abelian von Neumann algebra. Such a generalization, apart from its intrinsic mathematical interest, was then shown to be useful in discussions of the measurement process in quantum mechanics [NU62, Ume54]. Later, Davies and Lewis, in developing an operational basis for quantum measurement theory and a related theory of quantum stochastic processes, [DL70, Dav70, Dav76] gave a definition of a conditional expectation, as the dual to the concept of a measurement. This had the virtue of extending the von Neumann collapse postulate to observables with continuous spectra. Each measurement process then singled out an observable, defined as a POVM.

There is at least one more significant and independent line of development involving POVMs, which started in the early 1970s, and most notably expounded and developed in the works of Holevo and Helstrom [Hel76, Hol82]. Here the notion of a POVM appeared in the then emerging field of quantum information and statistical decision theory, essentially independently of the ideas of Ludwig on quantum foundations and Davies and Lewis on repeated measurements. This pioneering work led to the recognition of the fact that the statistics of optimal quantum measurements is often obtained by using what were called *non-orthogonal POVMs*. Simultaneously, the relevance of Naimark's extension theorem (from POVMs to projection valued measures), was first pointed out in this context.

The idea that POVMs could in some sense describe unsharp measurements was already present in the work of Davies and Lewis. These ideas were further developed to a more complete theory of fuzzy observables in [AE74, Pru84]. In the succeeding years the connection between unsharp or fuzzy observables and POVMs has been worked out in great detail by a number of authors. For a systematic presentation of various aspects of POVMs, we refer to [BGL97].

3. POSITIVE OPERATOR VALUED MEASURES

Quantum observables are mathematically described by positive operator valued measures. In this section we shortly recall this concept and fix the notation.

Let Ω be a Hausdorff locally compact second countable topological space (lcsc space, for short). We use the following notations:

- $\mathcal{B}(\Omega)$ is the Borel σ-algebra of Ω
- $C_0(\Omega)$ is the Banach space of countinuous complex functions on Ω vanishing at infinity, endowed with the uniform norm
- $M(\Omega)$ is the Banach space of complex measures on Ω endowed with the total variation norm
- $M(\Omega)^+$ is the subset of positive elements in $M(\Omega)$
- $P(\Omega) \subset M(\Omega)^+$ is the subset of probability measures

We will regard an element $\mu \in M(\Omega)$ both as a σ-additive mapping $\mu : \mathcal{B}(\Omega) \to \mathbb{C}$ and as a bounded linear functional $\mu : C_0(\Omega) \to \mathbb{C}$. In the latter case, we have the isometric isomorphism $M(\Omega) = C_0(\Omega)^*$. In general, if $L^\infty(\Omega)$ denotes the Banach space of bounded measurable complex functions on Ω endowed with the uniform norm, an element $\mu \in M(\Omega)$ determines a bounded linear functional $\mu : L^\infty(\Omega) \to \mathbb{C}$ by means of the formula

$$\mu(\phi) = \int_\Omega \phi(x)\, d\mu(x) \quad \forall \phi \in L^\infty(\Omega).$$

We let $B(\Omega)$ be the unit ball in $M(\Omega)$, and $B(\Omega)$ is endowed with the weak*-topology. Since $C_0(\Omega)$ is separable, $B(\Omega)$ is a compact metrizable space. Moreover, $B(\Omega)^+ := M(\Omega)^+ \cap B(\Omega)$ is closed in $B(\Omega)$, and $P(\Omega)$ is closed in $B(\Omega)^+$ if and only if Ω is compact.

Let $\mathcal{H}$ be a complex separable Hilbert space. We denote by $\mathcal{L}(\mathcal{H})$ the Banach space of bounded operators on $\mathcal{H}$ with the uniform norm. The convex set of positive trace one elements in $\mathcal{L}(\mathcal{H})$ is denoted by $\mathcal{S}(\mathcal{H})$.

A *positive operator valued measure* (POVM) E on Ω with values in $\mathcal{H}$ can be defined in two equivalent ways:

(a) as a mapping $\mathsf{E} : \mathcal{B}(\Omega) \to \mathcal{L}(\mathcal{H})$ such that
  (i) $\mathsf{E}(X) \geq O$ for all $X \in \mathcal{B}(\Omega)$
  (ii) $\mathsf{E}(\Omega) = I$
  (iii) for all disjoint finite or denumerable sequences $\{X_i\}_{i \in I}$ of sets in $\mathcal{B}(\Omega)$,

$$\mathsf{E}\left(\cup_{i \in I} X_i\right) = \sum\nolimits_{i \in I} \mathsf{E}(X_i),$$

  where the sum converges in the weak (or, equivalently, ultraweak or strong) operator topology;
(b) as a linear mapping $\mathsf{E} : C_0(\Omega) \to \mathcal{L}(\mathcal{H})$ such that
  (i) $\mathsf{E}(f) \geq O$ if $f \geq 0$
  (ii) If $\{f_n\}_{n \in \mathbb{N}}$ is a sequence of positive functions in $C_0(\Omega)$ such that $f_n \uparrow 1$, then $\mathsf{E}(f_n) \uparrow I$.

The link between the two definitions is established by the following integral formula

(1)
$$\mathsf{E}(f) = \int_\Omega f(x)\, d\mathsf{E}(x),$$

true for all $f \in C_0(\Omega)$. Here in the left hand side E is defined according to (b), and in the right hand side definition (a) is used. The integral in eq. (1) has to be understood in the following way. If E is a POVM in the sense of definition (a), then for all $T \in \mathcal{S}(\mathcal{H})$ we can define a probability measure $p_T^{\mathsf{E}} \in P(\Omega)$, given by

$$(2) \qquad\qquad p_T^{\mathsf{E}}(X) = \mathrm{tr}\,[T\mathsf{E}(X)] \quad \forall X \in \mathcal{B}(\Omega).$$

Eq. (1) then reads

$$\mathrm{tr}\,[T\mathsf{E}(f)] = \int_\Omega f(x)\,dp_T^{\mathsf{E}}(x) \quad \forall T \in \mathcal{S}(\mathcal{H}).$$

Note that if $\phi \in L^\infty(\Omega)$, we can define in the same way the bounded operator

$$(3) \qquad\qquad \mathsf{E}(\phi) = \int_\Omega \phi(x)\,d\mathsf{E}(x).$$

We see from eq. (3) that $\|\mathsf{E}(\phi)\| \le \|\phi\|$ for every $\phi \in L^\infty(\Omega)$. In particular, $\mathsf{E} : L^\infty(\Omega) \to \mathcal{L}(\mathcal{H})$ is continuous. We refer to [Ber66] for a more detailed discussion of these facts.

In quantum formalism, the set $\mathcal{S}(\mathcal{H})$ represents the *state space* of a quantum system. Quantum observables are, on the other hand, described by positive operator valued measures. The basic numerical predictions of quantum mechanics - the measurement outcome probability distributions - are given by the trace formula (2). Namely, the number $p_T^{\mathsf{E}}(X)$ is interpreted as the probability that a measurement outcome $x \in X$ occurs when a measurement of the observable E is performed and the system is in the state T.

Remark 1. If Ω is compact, then $C_0(\Omega) = C(\Omega)$, the space of continuous functions on Ω. Item (ii) in definition (b) is equivalent to $\mathsf{E}(1_\Omega) = I$, where 1_Ω is the constant function $1_\Omega(x) \equiv 1$.

For a subset $\mathcal{M} \subset \mathcal{L}(\mathcal{H})$, we denote by $\mathcal{M}'$ the commutant of $\mathcal{M}$ in $\mathcal{L}(\mathcal{H})$. If E is a POVM, we denote

$$(4) \qquad\qquad \mathsf{E}' = \{\mathsf{E}(f) \mid f \in C_0(\Omega)\}'.$$

Proposition 1. *Let E be a POVM. Then $\mathsf{E}' = \{\mathsf{E}(X) \mid X \in \mathcal{B}(\Omega)\}'$*

Proof. If U is unitary, then $U \in \mathsf{E}'$ if and only if the POVMs $f \mapsto \mathsf{E}(f)$ and $f \mapsto \mathsf{E}_U(f) = U\mathsf{E}(f)U^*$ are equal. This implies that $\mathsf{E}(X) = \mathsf{E}_U(X) = U\mathsf{E}(X)U^*$ for all $X \in \mathcal{B}(\Omega)$. Therefore, $U \in \mathsf{E}'$ if and only if $U \in \{\mathsf{E}(X) \mid X \in \mathcal{B}(\Omega)\}'$. The claim then follows by decomposing each $A \in \mathcal{L}(\mathcal{H})$ as a linear combination of four unitaries $\{U_1 \dots U_4\}$ such that $A' = \{U_1 \dots U_4\}'$. $\square$

4. Commutative observables

To emphasize the underlying physical context, we will use the term observable instead of POVM. This section, however, is concerned with certain mathematical representations of observables. The physical interpretation of these results are then discussed in Section 5.

Definition 1. An observable E is *commutative* if $\mathsf{E}(X) \in \mathsf{E}'$ for all $X \in \mathcal{B}(\Omega)$, or, equivalently, if $\mathsf{E}(f) \in \mathsf{E}'$ for all $f \in C_0(\Omega)$.

A special case of a commutative observable is a *projection valued measure* (PVM). It is a mapping $\mathsf{P} : \mathcal{B}(\Omega) \to \mathcal{L}(\mathcal{H})$ which in definition (a) satisfies the additional property

$$\mathsf{P}(X)\mathsf{P}(Y) = \mathsf{P}(X \cap Y)$$

for all $X, Y \in \mathcal{B}(\Omega)$. It is clear that commutativity of P follows from this condition. When using definition (b) this additional condition reads

$$\mathsf{P}(f)\mathsf{P}(g) = \mathsf{P}(fg)$$

for all $f, g \in C_0(\Omega)$. Projection valued measures are commonly referred as *sharp observables.*

As we show in this section, each commutative observable has at least two possible representations in terms of sharp observables. This topic has been recently studied by Jenčová and Pulmannová in [JP07],[JP09], where related results have been obtained. Our first representation theorem for commutative observables (Theorem 2) goes back to Holevo [Hol72], who proved it in the general case in which Ω is an arbitrary measurable space. It claims that, if E is a commutative observable, then there is a *unique* sharp observable P on the set of measures $B(\Omega)^+$ such that a kind of *canonical* spectral decomposition of E holds (eqs. (8) and (9)). Actually, P vanishes outside the set of probability measures $P(\Omega)$. As in the usual spectral theorem for normal operators, E and P have the same commutant in $\mathcal{L}(\mathcal{H})$. Under our assumption that Ω is a lcsc space, we will work out a different and more easy proof of Holevo's result, which simply makes use of elementary facts from C^*-algebra theory. A second representation theorem (Theorem 3) which, as explained below, is in a way complementary to the first, was proved in [Ali82]. We present it here without proof, describing only its content.

Theorem 1. *Let* $\mathsf{E} : C_0(\Omega) \to \mathcal{L}(\mathcal{H})$ *be a commutative observable. Then there exists a triple* $(\Sigma, \mathsf{P}, \mu)$ *such that*

(i) Σ *is a compact metrizable space;*
(ii) P *is a sharp observable (i.e. PVM) on* Σ *with values in* $\mathcal{H}$ *and* $\operatorname{supp} \mathsf{P} = \Sigma$;
(iii) $\mu : \Sigma \to B(\Omega)^+$ *is a weak*-continuous injective mapping;*
(iv) *for all* $f \in C_0(\Omega)$

$$(5) \qquad \mathsf{E}(f) = \int_\Sigma [\mu(h)](f) \, d\mathsf{P}(h).$$

If $(\tilde{\Sigma}, \tilde{\mathsf{P}}, \tilde{\mu})$ *is another triple satisfying items* (i)-(iv) *above, then there exists a homeomorphism* $\Phi : \Sigma \to \tilde{\Sigma}$ *such that* $\tilde{\mathsf{P}} = \mathsf{P} \circ \Phi^{-1}$ *and* $\mu = \tilde{\mu} \circ \Phi$.

Moreover, we have the equality $\mathsf{P}' = \mathsf{E}'$.

Proof. We denote by $C^*(\mathsf{E})$ the commutative C^*-algebra generated in $\mathcal{L}(\mathcal{H})$ by the set $\{\mathsf{E}(f) \mid f \in C_0(\Omega)\}$ and I. Since $C_0(\Omega)$ is separable and the mapping $f \mapsto \mathsf{E}(f)$ is continuous, the Banach space $C^*(\mathsf{E})$ is separable.

Let Σ be the maximal ideal space of $C^*(\mathsf{E})$. Then Σ is a compact metrizable space. We denote the Gelfand transform of $C^*(\mathsf{E})$ by Γ, i.e.,

$$\Gamma : C^*(\mathsf{E}) \to C(\Sigma), \qquad \Gamma A(h) = h(A) \qquad \forall A \in C^*(\mathsf{E}), \, h \in \Sigma \,.$$

Thus, Γ is a C^*-algebra isomorphism between $C^*(\mathsf{E})$ and $C(\Sigma)$.

For all $h \in \Sigma$, we define a mapping $\mu(h)$ from $C_0(\Omega)$ to $\mathbb{C}$ by

$$[\mu(h)](f) := [\Gamma\mathsf{E}(f)](h) \quad \forall f \in C_0(\Omega) \,.$$

From the properties of Γ and E follows that $\mu(h)$ is a positive bounded linear functional on $C_0(\Omega)$, hence $\mu(h) \in M(\Omega)^+$. Moreover,

$$|[\mu(h)](f)| \le \|\Gamma\| \, \|\mathsf{E}(f)\| \le \|f\| \,,$$

which implies that $\|\mu(h)\| \le 1$.

We have thus defined a mapping μ from Σ to $M(\Omega)$. The weak*-continuity of μ is clear from the definition. Since the subset $\{[\mu(\cdot)](f) \mid f \in C_0(\Omega)\}$ together with the constant function 1_Ω generates the algebra $C(\Sigma)$, it separates the points of Σ. Therefore, $\mu : \Sigma \to M(\Omega)$ is injective.

Since Γ^{-1} is a $*$-representation of $C(\Sigma)$ in $\mathcal{L}(\mathcal{H})$, there exists a sharp observable (PVM) $\mathsf{P} : \mathcal{B}(\Sigma) \to \mathcal{L}(\mathcal{H})$ such that

$$\Gamma^{-1}(\phi) = \int_\Sigma \phi(h) \, d\mathsf{P}(h) \quad \forall \phi \in C(\Sigma)$$

(see e.g. Theorem 2.5.5 in [Mur90]). Since Γ^{-1} is injective, $\operatorname{supp} \mathsf{P} = \Sigma$. Moreover,

$$\mathsf{E}(f) = \Gamma^{-1}\Gamma\mathsf{E}(f) = \int_\Sigma [\mu(h)](f) \, d\mathsf{P}(h) \,,$$

and item (iv) follows.

Now suppose $(\tilde{\Sigma}, \tilde{\mathsf{P}}, \tilde{\mu})$ is another triple satisfying items (i)-(iv) in the theorem. Let

$$\Pi(\tilde{\phi}) = \int_{\tilde{\Sigma}} \tilde{\phi}(\tilde{h}) \, d\tilde{\mathsf{P}}(\tilde{h}) \quad \forall \tilde{\phi} \in C(\tilde{\Sigma}) \,.$$

Then Π is a $*$-homomorphism of $C(\tilde{\Sigma})$ in $\mathcal{L}(\mathcal{H})$, which is injective since $\operatorname{supp} \tilde{\mathsf{P}} = \tilde{\Sigma}$. Moreover,

$$\Pi\left([\tilde{\mu}(\cdot)](f)\right) = \mathsf{E}(f) \quad \forall f \in C_0(\Omega).$$

By injectivity of $\tilde{\mu}$, the set of functions $\{[\tilde{\mu}(\cdot)](f) \mid f \in C_0(\Omega)\}$ separates the points of $\tilde{\Sigma}$, and it is closed under complex conjugation since $\tilde{\mu}(\tilde{h}) \in M(\Omega)^+$ for all $\tilde{h} \in \tilde{\Sigma}$. The functions $\{[\tilde{\mu}(\cdot)](f) \mid f \in C_0(\Omega)\}$ and 1_Ω thus generate the algebra $C(\tilde{\Sigma})$ by Stone-Weierstrass theorem. It follows that $\Pi(C(\tilde{\Sigma})) = C^*(E)$.

The composition $\Gamma\Pi$ is thus a $*$-isomorphism of $C(\tilde{\Sigma})$ into $C(\Sigma)$, hence there exists a homeomorphism $\Phi : \Sigma \to \tilde{\Sigma}$ such that

$$\Gamma\Pi(\tilde{\phi}) = \tilde{\phi} \circ \Phi \quad \forall \tilde{\phi} \in C(\tilde{\Sigma})$$

(see Theorem 2.1 in §VI of [Con90]). We thus have, for all $f \in C_0(\Omega)$,

$$[\mu(\cdot)](f) = \Gamma\Pi\left([\tilde{\mu}(\cdot)](f)\right) = [\tilde{\mu} \circ \Phi(\cdot)](f) \,.$$

Moreover,

$$\int_{\tilde{\Sigma}} \tilde{\phi}(\tilde{h}) \, d\tilde{\mathsf{P}}(\tilde{h}) = \Pi(\tilde{\phi}) = \Gamma^{-1}\left[\Gamma\Pi(\tilde{\phi})\right] = \int_\Sigma \tilde{\phi}(\Phi(h)) \, d\mathsf{P}(h) = \int_{\tilde{\Sigma}} \tilde{\phi}(\tilde{h}) \, d(\mathsf{P} \circ \Phi^{-1})(\tilde{h})$$

for all $\tilde{\phi} \in C(\tilde{\Sigma})$, which implies $\tilde{\mathsf{P}} = \mathsf{P} \circ \Phi^{-1}$.

Finally, $S \in \mathsf{E}'$ if and only if $S \in C^*(\mathsf{E})'$, and, by spectral theory, $C^*(\mathsf{E})' = \mathsf{P}'$, so that the last claim is proved. $\qquad\square$

Proposition 2. *Suppose that Ω is compact. Then in the situation of Theorem 1, $\mu(h)$ is a probability measure for all $h \in \Sigma$.*

Proof. For each $h \in \Sigma$, we get

$$[\mu(h)](1_\Omega) = [\Gamma\mathsf{E}(1_\Omega)](h) = \Gamma I(h) = h(I) = 1 \,.$$

$\qquad\square$

Suppose Σ is a lcsc space. We say that a mapping $\mu : \Sigma \to M(\Omega)^+$ is *weak*-measurable* if the mappings

$$\Sigma \ni h \to [\mu(h)](f) \in \mathbb{C}$$

are measurable for all $f \in C_0(\Omega)$. The following auxiliary result is needed later.

Proposition 3. *Let Σ be a lcsc space and μ a mapping from Σ to $M(\Omega)^+$. Then the following are equivalent:*

(a) *μ is weak*-measurable;*
(b) *the mappings $\Sigma \ni h \to [\mu(h)](X) \in \mathbb{R}$ are measurable for all $X \in \mathcal{B}(\Omega)$.*

If $\mu(\Sigma) \subset B(\Omega)^+$, then the above two conditions are also equivalent to

(c) *$\mu : \Sigma \to B(\Omega)^+$ is a measurable map.*

Proof. Suppose (b) holds. Fix $f \in C_0(\Omega)$. For $n \in \mathbb{N}$, let $\{E_i\}_{i \in I}$ be a finite partition of Ω into disjoint Borel sets such that $|f(x) - f(y)| \le 1/n$ for all $x, y \in E_i$ and $i \in I$. Choose $x_i \in E_i$, and let $\phi_n = \sum_i f(x_i) 1_{E_i}$. Then the mapping $h \mapsto [\mu(h)](\phi_n)$ is measurable, and

$$|[\mu(h)](f) - [\mu(h)](\phi_n)| \le \|\mu(h)\|/n \,.$$

Hence the mapping $h \mapsto [\mu(h)](f)$ is measurable being the pointwise limit of measurable functions.

Suppose then that (a) holds, i.e., μ is weak*-measurable. Let

$$L_\mu^\infty = \{\phi \in L^\infty(\Omega) \mid h \mapsto [\mu(h)](\phi) \text{ is a measurable mapping}\}\,.$$

If $\{\phi_n\}_{n \in \mathbb{N}} \subset L_\mu^\infty$, $\phi \in L^\infty(\Omega)$ are such that $\phi_n \uparrow \phi$, then $[\mu(h)](\phi_n) \uparrow [\mu(h)](\phi)$ for all h by dominated convergence theorem, hence $h \mapsto [\mu(h)](\phi)$ is measurable. This shows that L_μ^∞ is a monotone class in $L^\infty(\Omega)$. Since $C_0(\Omega) \subset L_\mu^\infty$, $L_\mu^\infty = L^\infty(\Omega)$ by Proposition 6.2.9 in [Ped89]. In particular, $h \mapsto [\mu(h)](X) = [\mu(h)](1_X)$ is measurable for all $X \in \mathcal{B}(\Omega)$.

Now assume that $\mu(\Sigma) \subset B(\Omega)^+$. Since the mapping $B(\Omega)^+ \ni \nu \mapsto \nu(f) \in \mathbb{C}$ is continuous, it is clear that (c) implies (a). On the other hand, suppose that (a) holds. Since $B(\Omega)^+$ is second countable, there exist sequences $\{\nu_n\}_{n \in \mathbb{N}}$ in $B(\Omega)^+$, $\{\varepsilon_n\}_{n \in \mathbb{N}}$ in $\mathbb{R}_+$, $\{k_n\}_{n \in \mathbb{N}}$ in $\mathbb{N}$, and, for all $n \in \mathbb{N}$, $\{f_i^n\}_{i=1}^{k_n}$ in $C_0(\Omega)$, such that the open sets

$$U_n = \{\nu \in B(\Omega)^+ \mid |\nu(f_i^n) - \nu_n(f_i^n)| < \varepsilon_n \text{ for all } i = 1, 2 \ldots k_n\}$$

form a denumerable basis for the topology of $B(\Omega)^+$. By (a), $\mu^{-1}(U_n) \in \mathcal{B}(\Sigma)$ for all n, hence $\mu^{-1}(U) \in \mathcal{B}(\Sigma)$ for all open $U \subset B(\Omega)^+$. This implies that μ is measurable. $\qquad\square$

Corollary 1. *The set $P(\Omega)$ of probability measures is a measurable subset in $B(\Omega)^+$.*

Proof. Take in the above proposition $\Sigma = B(\Omega)^+$ and μ the identity mapping, which is weak*-measurable being clearly weak*-continuous. The set $P(\Omega) = \{\nu \in B(\Omega)^+ \mid \nu(\Omega) = 1\}$ is thus measurable, since it is the inverse image of 1 under the measurable mapping $\mu \mapsto \mu(\Omega)$. $\square$

Proposition 4. *Let E and $(\Sigma, \mathsf{P}, \mu)$ be as in Theorem 1. Then $\mu(h)$ is a probability measure for P-almost all $h \in \Sigma$.*

Proof. If $\{f_n\}_{n \in \mathbb{N}}$ is a sequence of positive functions in $C_0(\Omega)$ such that $f_n \uparrow 1_\Omega$, we have

$$\int_\Sigma [\mu(h)](f_n) \, d\mathsf{P}(h) = \mathsf{E}(f_n) \uparrow I.$$

On the other hand, $[\mu(h)](f_n) \uparrow [\mu(h)](\Omega)$ for all $h \in \Sigma$, so that, by monotone convergence theorem and uniqueness of the limit,

$$(6) \qquad \int_\Sigma [\mu(h)](\Omega) \, d\mathsf{P}(h) = I.$$

For $\varepsilon > 0$, let $X_\varepsilon = \{h \in \Sigma \mid [\mu(h)](\Omega) < 1 - \varepsilon\}$. Multiplying eq. (6) by $\mathsf{P}(X_\varepsilon)$, we obtain

$$\mathsf{P}(X_\varepsilon) = \int_{X_\varepsilon} [\mu(h)](\Omega) \, d\mathsf{P}(h) \le (1 - \varepsilon)\mathsf{P}(X_\varepsilon),$$

hence $\mathsf{P}(X_\varepsilon) = 0$. The claim then follows. $\square$

Proposition 5. *Let E and $(\Sigma, \mathsf{P}, \mu)$ be as in Theorem 1. Then*

$$(7) \qquad \mathsf{E}(X) = \int_\Sigma [\mu(h)](X) \, d\mathsf{P}(h) \quad \forall X \in \mathcal{B}(\Omega).$$

Proof. For every $X \in \mathcal{B}(\Omega)$, define

$$\mathsf{F}(X) = \int_\Sigma [\mu(h)](X) \, d\mathsf{P}(h).$$

Then F is an observable (in the sense of definition (a)). If $C \subset \Omega$ is compact, and $\{f_n\}_{n \in \mathbb{N}}$ is a sequence in $C_0(\Omega)$ such that $f_n \downarrow 1_C$ (such a sequence exists by Lemma 6.2.8 in [Ped89]), then by dominated convergence theorem $\mathsf{E}(f_n) \downarrow \mathsf{E}(C)$, $[\mu(h)](f_n) \downarrow [\mu(h)](C)$ for all h, and $\int_\Sigma [\mu(h)](f_n) \, d\mathsf{P}(h) \downarrow \mathsf{F}(C)$. By eq. (5) and uniqueness of the limit, $\mathsf{E}(C) = \mathsf{F}(C)$.

As Ω is a lcsc space, every probability measure p_T^E, with $T \in \mathcal{S}(\mathcal{H})$, is regular (see e.g. Chapter 7 in [Fol99]). Therefore, the observable E has the following regularity property:

$$\mathsf{E}(X) = \text{LUB} \, \{\mathsf{E}(C) \mid C \text{ compact subset of } X\}.$$

It then follows that $\mathsf{E}(X) = \mathsf{F}(X)$ for all $X \in \mathcal{B}(\Omega)$. $\square$

If P is a PVM on $B(\Omega)^+$ such that $\mathsf{P}(B(\Omega)^+ \setminus P(\Omega)) = O$, then it is easy to check that the formula

$$\mathsf{E}(f) = \int_{B(\Omega)^+} \mu(f) \, d\mathsf{P}(\mu) \quad \forall f \in C_0(\Omega)$$

defines a commutative observable E on Ω. Collecting the above results, we obtain the following converse of this fact.

Theorem 2. *Let E be a commutative observable on Ω. Then there exists a unique sharp observable P on $B(\Omega)^+$ such that*

$$
\mathsf{E}(f) = \int_{B(\Omega)^+} \mu(f) \, d\mathsf{P}(\mu) \quad \forall f \in C_0(\Omega) , \tag{8}
$$

or, equivalently,

$$
\mathsf{E}(X) = \int_{B(\Omega)^+} \mu(X) \, d\mathsf{P}(\mu) \quad \forall X \in \mathcal{B}(\Omega) . \tag{9}
$$

Moreover, $B(\Omega)^+ \setminus P(\Omega)$ is a P-null subset of $B(\Omega)^+$, and $\mathsf{E}' = \mathsf{P}'$.

Proof. Apply Theorem 1, and observe that $\mu : \Sigma \to B(\Omega)^+$, being a continuous injective mapping between compact metric spaces, is a homeomorphism of Σ onto $\mu(\Sigma)$. Eq. (8) then follows transporting P on $\mu(\Sigma)$ by means of such homeomorphism and extending P to the whole $B(\Omega)^+$ by letting $\mathsf{P} = O$ on $B(\Omega)^+ \setminus \mu(\Sigma)$. Eq. (9) follows by Proposition 5. By Theorem 1, $\mathsf{P}' = \mathsf{E}'$. By Corollary 1 and Proposition 4, $B(\Omega)^+ \setminus P(\Omega)$ is a P-null Borel set.

Suppose P_1, P_2 are two PVMs on $B(\Omega)^+$ satisfying eq. (8). For $i = 1, 2$, let $\Sigma_i = \operatorname{supp}\mathsf{P}_i$, and let $j_i : \Sigma_i \to B(\Omega)^+$ be the inclusion mapping. Therefore,

$$
\mathsf{E}(f) = \int_{\Sigma_i} [j_i(\mu)](f) \, d\mathsf{P}_i(\mu) \quad \forall f \in C_0(\Omega).
$$

By Theorem 1, there exists a homeomorphism $\Phi : \Sigma_1 \to \Sigma_2$ such that $j_1 = j_2 \circ \Phi$ and $\mathsf{P}_2 = \mathsf{P}_1 \circ \Phi^{-1}$. Hence, $\Sigma_1 = \Sigma_2$, $\Phi = \operatorname{id}$, and $\mathsf{P}_1 = \mathsf{P}_2$. This shows uniqueness of P in eq. (8).

Suppose P_1, P_2 are two PVMs on $B(\Omega)^+$ satisfying eq. (9). If $f \in C_0(\Omega)$, choose a sequence $\{\phi_n\}_{n \in \mathbb{N}} \subset L^\infty(\Omega)$ as in the first half of the proof of Proposition 3. For all $\mu \in B(\Omega)^+$, we have $\mu(\phi_n) \to \mu(f)$, $|\mu(\phi_n)| \leq \|\phi_n\| \leq \|f\|$, and

$$
\int_{B(\Omega)^+} \mu(\phi_n) \, d\mathsf{P}_1(\mu) = \int_{B(\Omega)^+} \mu(\phi_n) \, d\mathsf{P}_2(\mu).
$$

By dominated convergence theorem and uniqueness of the limit,

$$
\int_{B(\Omega)^+} \mu(f) \, d\mathsf{P}_1(\mu) = \int_{B(\Omega)^+} \mu(f) \, d\mathsf{P}_2(\mu).
$$

Since this holds for all $f \in C_0(\Omega)$, $\mathsf{P}_1 = \mathsf{P}_2$ follows. $\qquad\square$

Theorem 2 can be thought of as expressing the commutative observable E as an *operator average* over probability measures. There is still another representation of a commutative observable, which we now discuss, and which in a sense expresses E as a probability average over sharp observables (PVMs) and is thus complementary to the preceding representation. A detailed proof of this result can be found in [Ali82, Ali84]. Here we only describe the representation.

Suppose E is an observable (not necessarily commutative), and consider the set E' defined in eq. (4). The double commutant $(\mathsf{E}')'$, which we denote by $\mathfrak{A}(\mathsf{E})$, is then a von Neumann algebra, and hence a weakly closed algebra of operators in $\mathcal{L}(\mathcal{H})$. (Recall that a von Neumann algebra $\mathfrak{A}$ is a $*$-invariant set of bounded operators on $\mathcal{H}$ which is an algebra under the operator product and equal to its double commutant, $\mathfrak{A} = \mathfrak{A}''$. Such an algebra is necessarily weakly closed). If $C^*(\mathsf{E})$ is the C^* algebra generated in $\mathcal{L}(\mathcal{H})$ by the set $\{\mathsf{E}(f) \mid f \in C_0(\Omega)\}$ and I, then clearly $C^*(\mathsf{E}) \subset \mathfrak{A}(\mathsf{E})$.

Next, consider the set $\widehat{\mathfrak{S}}$ of all sharp observables on Ω the ranges of which lie in $\mathfrak{A}(\mathsf{E})$. Thus, a sharp observable P is an element of $\widehat{\mathfrak{S}}$ if and only if $\mathsf{P}(X) \in \mathfrak{A}(\mathsf{E})$ for all $X \in \mathcal{B}(\Omega)$. There is a natural structure of a measure space on $\widehat{\mathfrak{S}}$, so that one can define ordinary Borel measures on it. Let $\overline{\nu}$ be such a positive measure on $\widehat{\mathfrak{S}}$, with $\overline{\nu}(\widehat{\mathfrak{S}}) = 1$ (i.e., $\overline{\nu}$ is a probability measure on $\widehat{\mathfrak{S}}$). Then the operators $\overline{\mathsf{E}}(X), X \in \mathcal{B}(\Omega)$, defined through the integral representation

$$\overline{\mathsf{E}}(X) = \int_{\widehat{\mathfrak{S}}} \mathsf{P}(X)\, d\overline{\nu}(\mathsf{P}), \quad \text{or, symbolically,} \quad \overline{\mathsf{E}} = \int_{\widehat{\mathfrak{S}}} \mathsf{P}\, d\overline{\nu}(\mathsf{P})$$

generate an observable. The convergence of the above integral is in the weak topology of $\mathcal{L}(\mathcal{H})$, meaning that for arbitrary $\phi, \psi \in \mathcal{H}$,

$$\langle \phi | \overline{\mathsf{E}}(X)\psi \rangle = \int_{\widehat{\mathfrak{S}}} \langle \phi | \mathsf{P}(X)\psi \rangle\; d\overline{\nu}(\mathsf{P})\,.$$

The interesting fact about the above representation of the observable $\overline{\mathsf{E}}$ as a probability average over the sharp observables P, is that $\overline{\mathsf{E}}$ and $\overline{\nu}$ determine each other uniquely.

If E is commutative, there exists a unique probability measure $\overline{\nu} = \nu$ for which $\overline{\mathsf{E}} = \mathsf{E}$. This result is stated precisely below.

Theorem 3. *Let E be a commutative observable on Ω. There exists a unique probability Borel measure ν on $\widehat{\mathfrak{S}}$ such that*

(10)
$$\mathsf{E}(X) = \int_{\widehat{\mathfrak{S}}} \mathsf{P}(X)\, d\nu(\mathsf{P})\,.$$

the integral converging weakly. The measure ν is unaltered if $\mathfrak{A}(\mathsf{E})$ is replaced by any other von Neumann algebra which contains $\mathfrak{A}(\mathsf{E})$.

A simple example. The representation theorems for commutative observables can be illustrated by means of a simple example. Let $\mathcal{H} = L^2(\mathbb{R}, dx)$ and let ν be a fixed probability measure defined on $\mathcal{B}(\mathbb{R})$. On $\mathcal{H}$ we define the commutative observable $\mathsf{E}(X)$, $X \in \mathcal{B}(\mathbb{R})$, as the operator of multiplication by the function $x \mapsto 1_X * \nu(x)$, i.e., for any $\phi \in \mathcal{H}$

$$(\mathsf{E}(X)\phi)(x) = 1_X * \nu(x)\phi(x)\,, \quad \text{where} \quad 1_X * \nu(x) = \int_{\mathbb{R}} 1_X(x - y)\, d\nu(y)\,,$$

1_X being the characteristic function of the set X. Defining the sharp observable P_x for each $x \in \mathbb{R}$ and writing $\mathsf{P}_0 = \mathsf{P}$:

$$(\mathsf{P}_x(X)\phi)(y) = 1_X(y - x)\phi(y)\,, \qquad X \in \mathcal{B}(\mathbb{R})\,, \phi \in \mathcal{H},$$

we easily see that $\mathsf{E}(X)$ may be written in the form

$$(11) \qquad \mathsf{E}(X) = \int_{\mathbb{R}} \mathsf{P}_x(X) \, d\nu(x) \, ,$$

as a weak integral. This is the representation given in eq. (10). On the other hand, it is also clear that $\mathsf{E}(X)$ may alternatively be written in the form:

$$(\mathsf{E}(X)\phi)(x) = \overline{\nu}_x(X)\phi(x) \, ,$$

again as a weak integral and where $\overline{\nu}$ is the probability measure $\overline{\nu}_x(X) = \nu(x - X)$ and $\overline{\nu}_0 = \overline{\nu}$. Thus, considering $x \mapsto \overline{\nu}_x(X)$ as a multiplication operator, for each fixed $X \in \mathcal{B}(\mathbb{R})$, we get

$$(12) \qquad \mathsf{E}(X) = \int_{\mathbb{R}} \overline{\nu}_x(X) \, d\mathsf{P}(x) = \int_{B(\mathbb{R})^+} \mu(X) \, d\tilde{\mathsf{P}}(\mu) \, ,$$

where $\tilde{\mathsf{P}} = \mathsf{P} \circ \Phi^{-1}$ with $\Phi : \mathbb{R} \to B(\mathbb{R})^+$, $\Phi(x) = \overline{\nu}_x$. The last formula is the expression of E in the form (9). In (11) the commutative observable E is written as an integral over the sharp observables P_x, with respect to the probability measure ν carried by these sharp observables, while in (12), E is written as an integral over the probability measures ν_x, with respect to the sharp observable P, supported by the probability measures. If $f \in C_0(\mathbb{R})$ is a real function, we may think of it as a sharp classical observable on the value space $\mathbb{R}$ while $f * \nu$ is an averaged out version of f. Thus, since

$$\mathsf{E}(f) = \int_{\mathbb{R}} \mathsf{P}_x(f) \, d\nu(x) = \int_{\mathbb{R}} \overline{\nu}_x(f) \, d\mathsf{P}(x) \, ,$$

we may call the first representation an unsharp quantum operator corresponding to a sharp classical observable and the second representation as the sharp quantum operator of an unsharp classical observable.

5. Fuzzy observables

Let Ω_1 and Ω_2 be two lcsc spaces. Recall that a mapping $\mu : \Omega_2 \to P(\Omega_1)$ is a *Markov kernel* if $x \mapsto [\mu(x)](X)$ is measurable for every $X \in \mathcal{B}(\Omega_1)$. By Proposition 3, this is equivalent to weak*-measurability and measurability of μ. We will sometimes use the notation $\mu_x(X) \equiv [\mu(x)](X)$.

Definition 2. Let E and F be two observables on Ω_1 and Ω_2, respectively. We say that E is a *fuzzy observable with respect to* F, or *fuzzy version of* F, if there exists a Markov kernel $\mu : \Omega_2 \to P(\Omega_1)$ such that

$$(13) \qquad p_T^{\mathsf{E}}(X) = \int_{\Omega_2} [\mu(x)](X) \, dp_T^{\mathsf{F}}(x) \quad \forall X \in \mathcal{B}(\Omega_1), \, T \in \mathcal{S}(\mathcal{H}) \, .$$

The condition (13) can be written as

$$(14) \qquad \mathsf{E}(X) = \int_{\Omega_2} [\mu(x)](X) \, d\mathsf{F}(x) \quad \forall X \in \mathcal{B}(\Omega_1)$$

or, equivalently,

$$\mathsf{E}(f) = \int_{\Omega_2} [\mu(x)](f) \, d\mathsf{F}(x) \quad \forall f \in C_0(\Omega_1)$$

where, as usual, the integrals are understood in the weak sense.

Definition 2 fits well with the concept of a fuzzy set. Namely, the mapping $[\mu(\cdot)](X)$ is a fuzzy set for each $X \in \mathcal{B}(\Omega_1)$. We can see the observable E as a composite mapping

$$X \mapsto [\mu(\cdot)](X) \mapsto \int_{\Omega_2} [\mu(x)](X) \, d\mathsf{F}(x) = \mathsf{F}([\mu(\cdot)](X)) \,.$$

The first mapping 'fuzzifies' outcome sets and then the observable F is applied on these fuzzy sets.

It is clear from (14) that if F is a commutative observable, then also its fuzzy version E is commutative. In particular, all fuzzy versions of sharp observables are commutative. The results of Section 4 imply that also the converse is true, namely, all commutative observables are fuzzy versions of sharp observables.

Obviously, the representations for a commutative observable given in Section 4 are very close to the definition of a fuzzy version. If the set of outcomes Ω is compact, then by taking into account Proposition 2 we see that (7) in Proposition 5 gives E as a fuzzy version of P. To see that the general (i.e. Ω not necessarily compact) representation form (9) formally satisfies the properties of Definition 2, let E be a commutative observable on Ω. We then have the sharp observable P given in Theorem 2. To show that E is a fuzzy version of P, we need to define a corresponding Markov kernel μ from $B(\Omega)^+$ to $P(\Omega)$. Fix a probability measure $\nu_0 \in P(\Omega)$ and define μ as

$$\mu : \quad \nu \mapsto \nu \quad \text{if } \nu \in P(\Omega) \,; \quad \nu \mapsto \nu_0 \quad \text{if } \nu \in B(\Omega)^+ \setminus P(\Omega) \,.$$

Then μ is measurable and hence a Markov kernel. As stated in Theorem 2, the set $B(\Omega)^+ \setminus P(\Omega)$ is a P-null set. Hence, we can write (9) in the form

$$\mathsf{E}(X) = \int_{B(\Omega)^+} \nu(X) \, d\mathsf{P}(\nu) = \int_{B(\Omega)^+} [\mu(\nu)](X) \, d\mathsf{P}(\nu) \,.$$

Therefore, E is a fuzzy version of P.

Example 1. Let us consider again the example discussed in the end of Section 4. Hence, let ν be a fixed probability measure on $\mathbb{R}$. It defines a Markov kernel

$$\mathbb{R} \ni x \mapsto \overline{\nu}_x \in P(\mathbb{R}) \,.$$

Here $\overline{\nu}_x$ is the probability measure $\overline{\nu}_x(X) = \nu(x - X)$. The fuzzification of a set $X \in \mathcal{B}(\mathbb{R})$ is now of the convolution form

$$X \mapsto [1_X * \nu](\cdot) \,.$$

This type of Markov kernels can be defined in a similar way when Ω is some other lcsc group than $\mathbb{R}$. They typically arise in the situations where the observables under investigation have a covariance property. We refer to [HLY04] for further details.

Being a fuzzy observable with respect to some other observable is a relative notion - one has to specify two observables that we are comparing together. We can also say that E is a *fuzzy observable* without further specification if it is a fuzzy version of some other observable F and this observable F is not a fuzzy version of E.

There are then also those observables which are not fuzzy observables. The following result has been proven by Jenčová and Pulmannová in [JP07].

Proposition 6. *A sharp observable P is not a fuzzy observable if and only if the von Neumann algebra $\mathfrak{A}(\mathsf{P})$ is a maximal commutative von Neumann subalgebra of $\mathcal{L}(\mathcal{H})$.*

This shows, in particular, that the usual sharp observables describing physical quantities such as position, momentum and spin of a particle and photon number of an electromagnetic field are not fuzzy observables. We will give an other formulation of this property in Section 6.

It is interesting to note that not all physical quantities are described by sharp observables and hence Proposition 6 cannot be applied to those cases. For instance, it would be interesting to know whether some of the phase observables are fuzzy or not. (See Chapter III in [BGL97] for an explanation of the phase observables and some other physical quantities not described by sharp observables.)

Markov kernels can add different kind of fuzziness to outcome sets. First of all, a Markov kernel can make outcome sets blurred. This happens, for instance, if in the situation of Example 1 the probability measure ν is of the form $d\nu(x) = f(x)dx$ and f is some smooth function (e.g. Gaussian function). This can typically model noise or imprecision in a measurement.

There is also another class of Markov kernels, which describe different kind of reduction. Namely, it can happen that a Markov kernel maps sets into sets (and not into genuine fuzzy sets), in which case there is thus no blurring. The possible reduction is now a consequence of the fact that two or more outcomes may become identified as one.

This latter kind of situation corresponds to a Markov kernel μ determined by a measurable function $\Phi : \Omega_2 \to \Omega_1$, i.e., $\mu(x) = \delta_{\Phi(x)}$ (δ_x is the Dirac measure at x). The function Φ simply relabels the outcomes, possibly giving same label to several different outcomes. Eq. (14) now reads

$$(15) \qquad\qquad \mathsf{E}(X) = \mathsf{F}(\Phi^{-1}(X)) \qquad \forall X \in \mathcal{B}(\Omega_1) \,,$$

which we also write as

$$\mathsf{E} = \mathsf{F} \circ \Phi^{-1} \,.$$

Our next proposition shows that fuzzification between two sharp observables can only be of the non-blurring form (15). This result follows immediately by combining Propositions 1 and 2 in [Hol85] even in the more general case in which P_2 is an observable (not necessarily sharp). In this general form it has also been proved in [JP07]. In our case, it is a simple application of Theorem 2. We present this new proof in order to illustrate the role of the uniqueness part of Theorem 2

Proposition 7. *Let P_1 and P_2 be sharp observables on Ω_1, Ω_2, respectively, with values in $\mathcal{H}$. If P_1 is a fuzzy version of P_2, then there exists a measurable function $\Phi : \Omega_2 \to \Omega_1$ such that $\mathsf{P}_1 = \mathsf{P}_2 \circ \Phi^{-1}$.*

Proof. Let $B_i^+ = B(\Omega_i)^+$. The map $\delta : \Omega_1 \to B_1^+$ given by

$$[\delta(x)](f) = f(x) \quad \forall x \in \Omega_1, f \in C_0(\Omega_1)$$

is a homeomorphism of Ω_1 into $\delta(\Omega_1)$, and $\delta(\Omega_1)$ is a locally closed subset of B_1^+.

Let $\mu : \Omega_2 \to B_1^+$ be a Markov kernel such that

(16)
$$\mathsf{P}_1(f) = \int_{\Omega_2} [\mu(x)](f) \, d\mathsf{P}_2(x) \quad \forall f \in C_0(\Omega_1) \,.$$

We claim that $\mu(x) \in \delta(\Omega_1)$ for P_2-almost all x. By eq. (16),

$$\mathsf{P}_1(f) = \int_{B_1^+} \nu(f) \, d(\mathsf{P}_2 \circ \mu^{-1})(\nu) \,.$$

On the other hand

$$\mathsf{P}_1(f) = \int_{\Omega_1} [\delta(x)](f) \, d\mathsf{P}_1(x) = \int_{B_1^+} \nu(f) \, d(\mathsf{P}_1 \circ \delta^{-1})(\nu) \,.$$

By the uniqueness part in Theorem 2, we have

$$\mathsf{P}_2 \circ \mu^{-1} = \mathsf{P}_1 \circ \delta^{-1} \,.$$

It follows that $\mu^{-1}(B_1^+ \setminus \delta(\Omega_1))$ is a P_2-null set, i.e., there exists $Z \in \mathcal{B}(\Omega_2)$ such that $\mathsf{P}_2(Z) = O$ and $\mu(x) \in \delta(\Omega_1)$ for all $x \in \Omega_2 \setminus Z$.

Setting $\Phi(x) = \delta^{-1}(\mu(x))$ for all $x \in \Omega_2 \setminus Z$ and $\Phi(x) = \text{constant}$ for all $x \in Z$, we obtain a measurable function $\Phi : \Omega_2 \to \Omega_1$, and

$$\begin{aligned}
(\mathsf{P}_2 \circ \Phi^{-1})(X) &= \mathsf{P}_2(\Phi^{-1}(X) \setminus Z) = \mathsf{P}_2(\mu^{-1}(\delta(X))) \\
&= \mathsf{P}_1(\delta^{-1}(\delta(X))) = \mathsf{P}_1(X) \,.
\end{aligned}$$

$\square$

6. JOINT MEASURABILITY

In quantum measurement theory, the traditional compatibility relation between two observables is *joint measurability*. If a pair of observables is jointly measurable, then their measurement outcome statistics can be recovered from the measurement statistics of a one single observable. The basic fact in quantum mechanics states that there are pairs of observables which are not jointly measurable. This opens up the room for many interesting questions.

The joint measurability relation and some related concepts are reviewed [Lah03]. Here we just recall the definition and recommend this reference for more explanation. As in earlier sections, we always suppose that Ω_1, Ω_2 are lcsc spaces.

Definition 3. Observables E_1 and E_2 are *jointly measurable* if there exists an observable G, defined on the outcome space $\Omega_1 \times \Omega_2$, such that

$$\mathsf{G}(X \times \Omega_2) = \mathsf{E}_1(X) \,, \qquad \mathsf{G}(\Omega_1 \times Y) = \mathsf{E}_2(Y)$$

for all $X \in \mathcal{B}(\Omega_1)$ and $Y \in \mathcal{B}(\Omega_2)$. In this case G is a *joint observable* of E_1 and E_2.

Having a one moment reflection on Definition 3, we notice that both E_1 and E_2 are special kind of fuzzy versions of G. Namely, if $\Phi_i : \Omega_1 \times \Omega_2 \to \Omega_i$ is the projection on the i-th factor, then $\mathsf{E}_i = \mathsf{G} \circ \Phi_i^{-1}$.

This observation leads to a natural generalization of joint measurability. Namely, instead of requiring that the observables E_1 and E_2 are marginals of some observable, we can be interested whether they are fuzzy versions of one single observable. If this would be the case and G is measured, we can calculate the measurement outcome distributions of E_1 and E_2. From this point of view, there is little difference whether E_1 and E_2 are marginals or just fuzzy versions of G.

It thus seems that the joint measurability requirement is too restrictive and one should instead concentrate on the compatibility relation related to fuzzy versions. However, this latter apparently more general relation is actually equivalent to joint measurability. This is the content of the following statement. The same result has been obtained in Proposition 3 of [Hol85] using a different method.

Proposition 8. *Two observables E_1 and E_2 are jointly measurable if and only if there exists an observable G such that both E_1 and E_2 are fuzzy versions of G.*

As we have discussed above, the 'only if' part of Proposition 8 is trivial since a marginal observable of G is its particular type of fuzzy version. Before giving the proof of Proposition 8 in the general situation, we demonstrate the equivalence of these two conditions in the case of finite outcome sets.

Let E_1 and E_2 be two observables with finite outcome sets $\Omega_1 = \{a_1, \ldots, a_m\}$ and $\Omega_2 = \{b_1, \ldots, b_n\}$, respectively. To simplify the notation, we denote $\mathsf{E}_1(\{a_i\}) \equiv \mathsf{E}_1(a_i)$, and so on. Suppose that E_1 and E_2 are both fuzzy versions of an observable G, which is defined on a finite outcome set Ω. This means that we have Markov kernels μ and ν such that

$$\mathsf{E}_1(a_i) = \sum_{x \in \Omega} \mu_x(a_i)\, \mathsf{G}(x)\,, \qquad \mathsf{E}_2(b_j) = \sum_{x \in \Omega} \nu_x(b_j)\, \mathsf{G}(x)\,.$$

For every $a_i \in \Omega_1, b_j \in \Omega_2, x \in \Omega$, denote

$$\lambda_x(a_i, b_j) = \mu_x(a_i)\nu_x(b_j)\,.$$

It is clear that $0 \leq \lambda_x(a_i, b_j) \leq 1$ and

$$\sum_{a_i, b_j} \lambda_x(a_i, b_j) = 1 \quad \forall x \in \Omega\,.$$

Thus, λ is a Markov kernel.

We now define an observable $\widetilde{\mathsf{G}}$ with the outcome set $\Omega_1 \times \Omega_2$ as

$$\widetilde{\mathsf{G}}(a_i, b_j) = \sum_{x \in \Omega} \lambda_x(a_i, b_j)\, \mathsf{G}(x)\,.$$

By its definition, $\widetilde{\mathsf{G}}$ is a fuzzy version of G. Moreover, a simple calculation gives

$$\sum_{a_i \in \Omega_1} \widetilde{\mathsf{G}}(a_i, b_j) = \sum_{x \in \Omega} \sum_{a_i \in \Omega_1} \lambda_x(a_i, b_j)\mathsf{G}(x) = \sum_{x \in \Omega} \nu_x(b_j)\mathsf{G}(x) = \mathsf{E}_2(b_j)$$

and similarly

$$\sum_{b_j \in \Omega_2} \widetilde{\mathsf{G}}(a_i, b_j) = \mathsf{E}_1(a_i)\,.$$

We conclude that $\widetilde{\mathsf{G}}$ is a joint observable for E_1 and E_2.

Before proving Proposition 8, we need a small technical lemma.

Lemma 1. *Let $\Sigma, \Omega_1, \Omega_2$ be lcsc spaces. Suppose $\mu : \Sigma \to P(\Omega_1)$, $\nu : \Sigma \to P(\Omega_2)$ are Markov kernels. Then the mapping*

$$\lambda : \Sigma \to P(\Omega_1 \times \Omega_2), \quad \lambda(h) = \mu(h) \otimes \nu(h)$$

is a Markov kernel.

Proof. If $f_1 \in C_0(\Omega_1)$, $f_2 \in C_0(\Omega_2)$, then the mapping

$$h \mapsto [\lambda(h)](f_1 \otimes f_2) = [\mu(h)](f_1)[\nu(h)](f_2)$$

is measurable by Proposition 3. If $g \in C_0(\Omega_1 \times \Omega_2)$, then there is a sequence $\{g_n\}_{n \in \mathbb{N}} \subset C_0(\Omega_1) \otimes C_0(\Omega_2)$ such that $g_n \to g$ in the uniform norm. It follows that $[\lambda(h)](g_n) \to [\lambda(h)](g)$ for all h, and thus the mapping $h \mapsto [\lambda(h)](g)$ is measurable. Another application of Proposition 3 then gives the claim. $\square$

Proof of Proposition 8. The proof in the general case follows the same lines as the finite outcome example above.

Suppose that Ω_1, Ω_2 are the outcome spaces for $\mathsf{E}_1, \mathsf{E}_2$, respectively. We assume that both E_1 and E_2 are fuzzy versions of the observable G, which is defined on the outcome space Ω. We need to show that there exists a joint observable $\widetilde{\mathsf{G}}$ of E_1 and E_2.

For every $X \in \mathcal{B}(\Omega_1), Y \in \mathcal{B}(\Omega_2), x \in \Omega$, denote

$$\lambda(x) = \mu(x) \otimes \nu(x)\,.$$

Then the mapping $x \mapsto \lambda(x)$ is a Markov kernel by Lemma 1.

We define an observable $\widetilde{\mathsf{G}}$ by formula

$$\widetilde{\mathsf{G}}(Z) = \int [\lambda(x)](Z)\, d\mathsf{G}(x) \quad Z \in \mathcal{B}(\Omega_1 \times \Omega_2).$$

It follows directly from our construction that

$$\widetilde{\mathsf{G}}(X \times \Omega_2) = \mathsf{E}_1(X)\,, \qquad \widetilde{\mathsf{G}}(\Omega_1 \times Y) = \mathsf{E}_2(Y)$$

for all $X \in \mathcal{B}(\Omega_1)$ and $Y \in \mathcal{B}(\Omega_2)$. Hence, $\widetilde{G}$ is a joint observable of E_1 and E_2. $\square$

In Section 5 we discussed about observables which are not fuzzy (in the absolute sense). This property has an equivalent formulation in terms of joint measurements, which is expressed in the next proposition.

Proposition 9. *Let* E *be an observable. The following conditions are equivalent:*

(a) E *is not a fuzzy observable.*

(b) *If* F *is another observable which is jointly measurable with* E*, then* F *is a fuzzy version of* E*.*

Proof. Suppose that (a) holds. Let F be an observable which is jointly measurable with E, and let G be their joint observable. This means, in particular, that E is a fuzzy version of G. By the assumption, this can happen only if G is a fuzzy version of E. The relation of "being fuzzy version" is transitive (see e.g. [Hei05]). Therefore, we conclude that F is a fuzzy version of E. Thus, (b) holds.

Suppose then that (a) does not hold, i.e., E is a fuzzy observable. Then there is an observable F such that E is a fuzzy version of F but F is not a fuzzy version of E. Since F is trivially fuzzy version of itself, we conclude from Proposition 8 that E and F are jointly measurable. Hence, (b) does not hold. □

References

[AE74]   S.T. Ali and G.G. Emch. Fuzzy observables in quantum mechanics. *J. Math. Phys.*, 15:176–182, 1974.

[Ali82]   S.T. Ali. A geometrical property of POV-measures and systems of covariance. In H.-D. Doebner, S.I. Andersson, and H.R. Petry, editors, *Differential geometric methods in mathematical physics*, volume 905 of *Lecture Notes in Mathematics*, pages 207–228, Berlin, 1982. Springer-Verlag.

[Ali84]   S.T. Ali. Commutative systems of covariance and a generalization of Mackey's imprimitivity theorem. *Canad. Math. Bull.*, 27:390–397, 1984.

[Ber66]   S.K. Berberian. *Notes on Spectral Theory*. D. Van Nostrand Company, Princeton, New Jersey, 1966.

[BGL97]   P. Busch, M. Grabowski, and P.J. Lahti. *Operational Quantum Physics*. Springer-Verlag, Berlin, 1997. second corrected printing.

[BHL07]   P. Busch, T. Heinonen, and P. Lahti. Heisenberg's uncertainty principle. *Phys. Rep.*, 452:155–176, 2007.

[Con90]   J.B. Conway. *A Course in Functional Analysis*. Springer, 2nd edition, 1990.

[Dav70]   E.B. Davies. On the repeated measurements of continuous observables in quantum mechanics. *J. Funct. Anal.*, 6:318–346, 1970.

[Dav76]   E.B. Davies. *Quantum Theory of Open Systems*. Academic Press, London, 1976.

[DL70]   E.B. Davies and J.T. Lewis. An operational approach to quantum probability. *Comm. Math. Phys.*, 17:239–260, 1970.

[Fol99]   G.B. Folland. *Real Analysis*. John Wiley & Sons Inc., New York, second edition, 1999.

[Hei05]   T. Heinonen. Optimal measurements in quantum mechanics. *Phys. Lett. A*, 346:77–86, 2005.

[HLY04]   T. Heinonen, P. Lahti, and K. Ylinen. Covariant fuzzy observables and coarse-graining. *Rep. Math. Phys.*, 53:425–441, 2004.

[Hel76]   C. W. Helstrom *Quantum Detection and Estimation Theory*. Academic Press, New York, 1976.

[Hol72]   A. S. Holevo. An analogue of statistical decision theory and noncommutative probability theory. *Trudy Moskov. Mat. Obšč.*, 26:133–149, 1972.

[Hol82]   A. S. Holevo *Probabilistic and Statistical Aspects of Quantum Theory*. North Holland, Amsterdam, 1982.

[Hol85]   A. S. Holevo. Statistical definition of observable and the structure of statistical models. *Rep. Math. Phys.*, 22:385–407, 1985.

[JP07]   A. Jenčová and S. Pulmannová. How sharp are PV measures? *Rep. Math. Phys.*, 59:257–266, 2007.

[JP09]   A. Jenčová and S. Pulmannová. Characterizations of commutative POV measures. *Found. Phys.*, 2009. Published online: 21 January 2009.

[Lah03] P. Lahti. Coexistence and joint measurability in quantum mechanics. *Int. J. Theor. Phys.*, 42:893–906, 2003.

[Lud64] G. Ludwig. Versuch einer axiomatischen Grundlegung der Quantenmechanik und allgemeinerer physikalischer Theorien. *Z. Physik*, 181:233–260, 1964.

[Lud70] G. Ludwig. *Deutung des Begriffs "physikalische Theorie" und axiomatische Grundlegung der Hilbertraumstruktur der Quantenmechanik durch Hauptsätze des Messens.* Springer-Verlag, Berlin, 1970.

[Moy54] S.-T. C. Moy. Characterizations of conditional expectations as a transformation on function spaces. *Pacific J. Math.*, 4:47–64, 1954.

[Mur90] G.J. Murphy. C^*-*algebras and Operator Theory*. Academic Press, Inc., San Diego, 1990.

[NT54]   M. Nakamura and T. Turumaru. Expectations in an operator algebra. *Tôhoku Math J.*, 6:182–188, 1954.

[NU62]   M. Nakamura and H. Umegaki. On von Neumann's theory of measurements in quantum statistics. *Math. Japon.*, 7:151–157, 1962.

[Ped89] G. Pedersen. *Analysis now.* Springer-Verlag, New York, 1989.

[Pru84] E. Prugovečki. *Stochastic Quantum Mechanics and Quantum Spacetime.* D. Reidel Publishing Co., Dordrecht, 1984.

[Ume54] H. Umegaki. Conditional expectation in an operator algebra. *Tôhoku Math. J.*, 6:177–181, 1954.

S. Twareque Ali, Department of Mathematics and Statistics, Concordia University, Montreal, Quebec, Canada H3G 1M8
   E-mail address: stali@mathstat.concordia.ca

Claudio Carmeli, Dipartimento di Fisica, Università di Genova and I.N.F.N., Sezione di Genova, Via Dodecaneso 33, 16146 Genova, Italy
   E-mail address: carmeli@ge.infn.it

Teiko Heinosaari, Niels Bohr Institute, Copenhagen University, Denmark and Department of Physics, University of Turku, Finland
   E-mail address: heinosaari@gmail.com

Alessandro Toigo, Dipartimento di Informatica, Università di Genova, Via Dodecaneso 35 and I.N.F.N., Sezione di Genova, Via Dodecaneso 33, 16146 Genova, Italy
   E-mail address: toigo@ge.infn.it