Found Phys (2009) 39: 613–624
DOI 10.1007/s10701-009-9273-1

Characterizations of Commutative POV Measures

Anna Jenčová · Sylvia Pulmannová

Received: 5 November 2008 / Accepted: 7 January 2009 / Published online: 21 January 2009

Abstract Two different characterizations of POV measures with commutative range are compared using a representation of some stochastic operators by (weak) Markov kernels. A representation by Choquet theorem is obtained as an integral over functions of a sharp observable appearing in one of the characterizations. A Naimark extension is constructed.

Keywords PV measures · POV measures · Observables · Markov kernel · Weak Markov kernel · Smearing · Choquet theorem · Naimark extension

1 Introduction

Quantum observables are usually considered as normalized POV measures. The special subclass represented by PV measures is classified as sharp observables, while the remaining POV measures are classified as unsharp. Every POV measure M is completely described by the family $\{m \circ M\}$ of probability distributions in all physical states m. This enables us to apply some results from classical probability theory and theory of experiments to quantum observables. For example, a stochastic operator is defined as an affine mapping from the set of probability measures on one measurable space to the set of probability measures on another measurable space. This yields the following relation between two observables M and N: we say that N is a smearing

This work was supported by Center of excellence SAS, CEPI I/2/2005 and grant APVV-0071-06 and grant VEGA 2/0032/09 SAS.

A. Jenčová · S. Pulmannová (✉)
Mathematical Institute, Slovak Academy of Sciences, Štefánikova 49, 814 73 Bratislava, Slovakia
e-mail: pulmann@mat.savba.sk

A. Jenčová
e-mail: jenca@mat.savba.sk

of M if there is a stochastic operator that maps $m \circ M$ to $m \circ N$, and write $M \preceq N$. The relation $\preceq$ is a preorder on the set of observables.

In this paper, we study the special case of POV measures with commutative range. Observables of this type were characterized in several different ways, but always as smearings of sharp observables (PV-measures).

The first characterization we consider is based on the von Neumann theorem stating that any system of mutually commuting bounded self-adjoint operators on a separable Hilbert space can be represented by functions of one bounded self-adjoint operator (cf. [11, Theorem 10], [7, Théorème p. 355]). In [5] (see Theorem A below), we proved that a POV measure with commutative range is a smearing of the PV-measure corresponding to the self-adjoint operator with the property that all elements in the range are its functions.

The second result was obtained by Holevo [3], where a characterization is given by a smearing of a PV measure defined on the Borel σ-algebra of subsets of a set of probability measures (see Theorem B below).

In addition, Ali [1] obtained an integral representation for (not necessarily commutative) POV measures defined on a separable metrizable locally compact topological space by means of Choquet theorem.

Using a representation of some stochastic operators by (weak) Markov kernels, we show equivalence of Theorems A and B. In fact, the sharp observable in Theorem B is a smearing of the sharp observable of Theorem A. Conversely, the sharp observable in Theorem A can be chosen as a smearing of the observable in Theorem B.

In addition, we show that a representation by Choquet theorem for a POV measure with commutative range defined on a separable metrizable locally compact topological space can be obtained as an integral over the functions of the sharp observable from Theorem A. Finally, Theorem A can also be used to the construction of a Naimark extension for the POV measure.

2 Stochastic Maps and Markov Kernels

Let $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ be measurable spaces and let $\mathcal{P}(X, \mathcal{A})$ be the set of probability measures on $(X, \mathcal{A})$. An affine map

$$T : \mathcal{P}(X, \mathcal{A}) \to \mathcal{P}(Y, \mathcal{B})$$

is called a stochastic map, or a stochastic operator. It is clear that a stochastic map extends to a positive linear mapping on the space of complex measures, preserving the total variation norm on the positive cone. This extended map will be also called a stochastic map.

An important class of stochastic maps is given by means of Markov kernels, that is, mappings $\lambda : X \times \mathcal{B} \to [0, 1]$, satisfying the following conditions:

(1) the function $x \mapsto \lambda(x, B)$ is $\mathcal{A}$-measurable for all fixed $B \in \mathcal{B}$
(2) $B \mapsto \lambda(x, B)$ is a probability measure on $\mathcal{B}$ for any fixed $x \in X$

It is easy to see that

$$B \mapsto \int_X \lambda(x, B)\rho(dx) \tag{1}$$

is a probability measure on $(Y, \mathcal{B})$ for any $\rho \in \mathcal{P}(X, \mathcal{A})$ and that this defines a stochastic map, which we denote by T_λ.

In general, there is little hope that a stochastic map is defined by a Markov kernel. The situation is different if we consider stochastic maps defined on certain convex subsets of $\mathcal{P}(X, \mathcal{A})$. Let μ be a σ-finite measure on $(X, \mathcal{A})$ and let $L(\mu) = \{\rho \in \mathcal{P}(X, \mathcal{A}),\ \rho \ll \mu\}$.

Theorem 2.1 [8, pp. 338–340] *Let μ be a σ-finite measure on $(X, \mathcal{A})$ and let $(Y, \mathcal{B})$ be a standard Borel space. Let $T : L(\mu) \to \mathcal{P}(Y, \mathcal{B})$ be a stochastic operator. Then there is a Markov kernel $\lambda : X \times \mathcal{B} \to [0, 1]$, such that $T = T_\lambda|_{L(\mu)}$.*

2.1 Weak Markov Kernels

Let $\tau \in \mathcal{P}(X, \mathcal{A})$. A function $\nu : X \times \mathcal{B} \to [0, 1]$ is called a *weak Markov kernel with respect to τ* (or a τ-Markov kernel) if it satisfies:

(i) $x \mapsto \nu(x, B)$ is $\mathcal{A}$-measurable for each $B \in \mathcal{B}$
(ii) $\nu(x, \emptyset) = 0$ and $\nu(x, Y) = 1$, τ-a.s.
(iii) for each sequence $B_n \in \mathcal{B}$, such that $B_n \cap B_m = \emptyset$ whenever $n \neq m$, we have $\nu(x, \bigcup_n B_n) = \sum_n \nu(x, B_n)$, τ-a.s.

It is easy to see that a weak Markov kernel ν defines a stochastic map $T_\nu : L(\tau) \to L(T_\nu(\tau))$, by (1). Therefore T_ν extends to a stochastic operator $T_\nu : L_1(X, \mathcal{A}, \tau) \to L_1(Y, \mathcal{B}, T_\nu(\tau))$. Such an operator is always bounded and its dual $T_\nu^* : L_\infty(Y, \mathcal{B}, T_\nu(\tau)) \to L_\infty(X, \mathcal{A}, \tau)$ is a Markov operator, that is, positive, linear operator with $T_\nu^*(1) = 1$ and $T_\nu^*(f_n) \downarrow 0$ whenever $f_n \downarrow 0$ (see, e.g., [9]).

Suppose that ν_1 and ν_2 are weak Markov kernels with respect to τ. Then it is clear that $T_{\nu_1} = T_{\nu_2}$ if and only if

$$\nu_1(x, B) = \nu_2(x, B), \quad \tau\text{-a.s.},\ B \in \mathcal{B} \tag{2}$$

This defines an equivalence relation $\sim_\tau$ on the set of τ-Markov kernels. Let us denote the quotient of this set by $\sim_\tau$ by $\mathcal{K}(X, \mathcal{B}, \tau)$.

It is clear that a Markov kernel is a weak Markov kernel with respect to any $\tau \in \mathcal{P}(X, \mathcal{A})$. On the other hand, if $(Y, \mathcal{B})$ is a standard Borel space and ν a weak Markov kernel, then Theorem 2.1 implies that there exists a Markov kernel λ, such that $T_\nu = T_\lambda|_{L(\tau)}$. This means that the equivalence class of ν with respect to $\sim_\tau$ contains a Markov kernel, in this case. Note that then

$$T_\lambda^*(f) = \int_Y f(y)\lambda(x, dy)$$

for $f \in L_\infty(Y, \mathcal{B}, T_\lambda(\tau))$.

Another important example of a weak Markov kernel is the conditional probability: let $X = Y$ and let $\mathcal{A} \subset \mathcal{B}$, then $\nu(x, B) := P(B|\mathcal{A})(x)$ defines a weak Markov kernel with respect to $\tau = P|_\mathcal{A}$.

2.2 Composition of Weak Markov Kernels

Let $(Z,\mathcal{C})$ be another measurable space and let $\mu : Y \times \mathcal{C} \to [0,1]$ be a weak Markov kernel with respect to some $\rho \in \mathcal{P}(Y,\mathcal{B})$. If we have $T_\nu(\tau) \ll \rho$, then $L(T_\nu(\tau)) \subseteq L(\rho)$, so that there exists a stochastic operator $T_\mu \circ T_\nu : L(\tau) \to \mathcal{P}(Z,\mathcal{C})$. Our aim is to show that this stochastic operator is given by a weak Markov kernel $\lambda : X \times \mathcal{C} \to [0,1]$, with respect to τ.

So let $\omega \in L(\tau)$ and $C \in \mathcal{C}$. Let us denote $\mu_C = \mu(\cdot, C)$. Then

$$T_\mu(T_\nu(\omega))(C) = \int_Y \mu(y,C) T_\nu(\omega)(dy) = \int_X T_\nu^*(\mu_C)(x)\omega(dx) \tag{3}$$

Let us put

$$\lambda(x,C) = T_\nu^*(\mu_C)(x) \tag{4}$$

and check that this defines a weak Markov kernel with respect to τ.

First, since $\lambda_C = T_\nu^*(\mu_C) \in L_\infty(X,\mathcal{A},\tau)$, it is obviously $\mathcal{A}$-measurable, so that condition (i) is satisfied. Further, by (3), we get $\int_X \lambda(x,\emptyset)\omega(dx) = T_\mu \circ T_\nu(\omega)(\emptyset) = 0$ and similarly $\int_X \lambda(x,Z)\omega(dx) = 1$ for all $\omega \ll \tau$, this implies condition (ii). Let now $\{C_n\}$ be a sequence of pairwise disjoint elements in $\mathcal{C}$. Then, again by (3), we get for each k that

$$\lambda\left(x, \bigcup_{n=1}^k C_n\right) = \sum_{n=1}^k \lambda(x,C_n), \qquad \tau\text{-a.s.}$$

Then $\sum_{n=1}^k \lambda(x,C_n)$ is τ-a.s. an increasing sequence of positive functions, bounded by 1. It follows that the limit $\sum_n \lambda(x,C_n)$ exists (τ-a.s.) and that

$$\begin{aligned}\int_X \sum_n \lambda(x,C_n)\omega(dx) &= \sum_n \int_X \lambda(x,C_n)\omega(dx) = \sum_n T_\mu \circ T_\nu(\omega)(C_n) \\ &= T_\mu \circ T_\nu(\omega)\left(\bigcup_n C_n\right) = \int_X \lambda\left(x, \bigcup_n C_n\right)\omega(dx)\end{aligned}$$

for all $\omega \ll \tau$, this entails (iii).

Note that if ν and μ are Markov kernels, then the expression

$$\lambda(x,C) = \int_Y \mu(y,C)\nu(x,dy)$$

defines a Markov kernel, such that $T_\lambda = T_\mu \circ T_\nu$ (in fact, it is enough if ν is a Markov kernel). We will use the expression $\lambda = \mu \circ \nu$ to denote the weak Markov kernel defined by (4).

3 Observables and their Smearings

Let $\mathcal{H}$ be a separable Hilbert space, $\mathcal{E}(\mathcal{H}) = \{T \in B(\mathcal{H}) : 0 \le T \le I\}$ the set of Hilbert space effects, and let m_0 be a faithful state on $B(\mathcal{H})$. Let $M : (X,\mathcal{A}) \to \mathcal{E}(\mathcal{H})$

be a POV measure (observable) and let $\nu : X \times \mathcal{B} \to [0,1]$ be a weak Markov kernel with respect to the measure $m_0 \circ M$. Then

$$B \to \int_X \nu(x, B) M(dx)$$

defines a POV measure $\nu \circ M : (Y, \mathcal{B}) \to \mathcal{E}(\mathcal{H})$, called the *smearing of M by ν*. Clearly, $\nu_1 \circ M = \nu_2 \circ M$ if and only if $\nu_1 \sim_{m_0 \circ M} \nu_2$. Note that if m is a state on $B(\mathcal{H})$, then

$$m \circ (\nu \circ M)(B) = \int_X \nu(x, B) m \circ M(dx) = T_\nu(m \circ M)(B)$$

If N is a smearing of M, we write $M \preceq N$. The relation $\preceq$ is clearly reflexive. We next show that it is also transitive. Indeed, let $M : (X, \mathcal{A}) \to \mathcal{E}(\mathcal{H})$, $N : (Y, \mathcal{B}) \to \mathcal{E}(\mathcal{H})$ and $R : (Z, \mathcal{C}) \to \mathcal{E}(\mathcal{H})$ be POV measures, such that $N = \nu \circ M$ and $R = \mu \circ N$, where $\nu : X \times \mathcal{B} \to [0,1]$ is an $m_0 \circ M$-Markov kernel and $\mu : Y \times \mathcal{C} \to [0,1]$ is an $m_0 \circ N$-Markov kernel. Since $m_0 \circ N = T_\nu(m_0 \circ M)$, we can define the $m_0 \circ M$-Markov kernel $\lambda = \mu \circ \nu : X \times \mathcal{C} \to [0,1]$, as in the previous section. Then for every state m,

$$m \circ (\lambda \circ M) = T_\lambda(m \circ M) = T_\mu(T_\nu(m \circ M)) = T_\mu(m \circ N) = m \circ R,$$

so that $R = \lambda \circ M$ and $M \preceq R$. Therefore, $\preceq$ is a preorder on the set of observables. If $M \preceq R$ and $R \preceq M$ we write $M \sim R$, and say that M and R are equivalent.

4 Observables with Commutative Range

Suppose that the range $\mathcal{R}(M)$ of the observable $M : (X, \mathcal{A}) \to \mathcal{E}(\mathcal{H})$ is commutative. Observables of this type were characterized in several different ways, but they are always obtained from PV measures (sharp observables).

An example of such observable is a POV measure $M : (X, \mathcal{A}) \to L_\infty(Y, \mathcal{B}, \tau) \subset B(L_2(Y, \mathcal{B}, \tau))$ for some $\tau \in \mathcal{P}(Y, \mathcal{B})$. It is easy to show that the set of all such POV measures can be identified with the quotient $\mathcal{K}(Y, \mathcal{A}, \tau)$, the correspondence is given by $M(A) = \nu_A$. Note that if $N = \mu \circ M$ for a weak Markov kernel $\mu : X \times \mathcal{C} \to [0,1]$, then N corresponds to the composition $\mu \circ \nu$.

In fact, this describes all observables with commutative range. Indeed, let $\mathcal{R}$ be any commutative von Neumann subalgebra in $B(\mathcal{H})$, containing the range of M. Then $\mathcal{R}$ is generated by a single self-adjoint operator S. Let P_S be the spectral measure of S and let us denote $\tau_S = m_0 \circ P_S$. Then $\mathcal{R}$ can be identified with the von Neumann algebra $L_\infty(\mathbb{R}, B(\mathbb{R}), \tau_S)$, by $f \mapsto f(S)$. It follows that M corresponds to a τ_S-Markov kernel $\nu : \mathbb{R} \times \mathcal{A} \to [0,1]$, so that

$$M(A) = \nu_A(S) = \int_{\mathbb{R}} \nu(t, A) P_S(dt) \tag{5}$$

This implies the following characterization of commutative POV measures.

Theorem A [5] *The observable M has a commutative range if and only if there is a PV measure $P : (\mathbb{R}, B(\mathbb{R})) \to \mathcal{E}(\mathcal{H})$ and a weak Markov kernel $\nu : \mathbb{R} \times \mathcal{A} \to [0,1]$, such that $M = \nu \circ P$.*

Our next aim is to compare this characterization to another one, obtained by Holevo [3]. For this, let $D := \mathcal{P}(X, \mathcal{A}) \subset [0,1]^{\mathcal{A}}$. Let $B([0,1]^{\mathcal{A}})$ be the Borel σ-algebra on $[0,1]^{\mathcal{A}}$, generated by the cylinders, and let $\mathcal{S} := B([0,1]^{\mathcal{A}}) \cap D$.

Theorem B *M is an observable with commutative range if and only if there exists a (unique) PV measure $\tilde{P} : (D, \mathcal{S}) \to \mathcal{E}(\mathcal{H})$, such that*

$$M(A) = \int_D \rho(A)\tilde{P}(d\rho)$$

Note that also in Theorem B, M is a smearing of the PV measure $\tilde{P}$ by the Markov kernel $\kappa : D \times \mathcal{A} \to [0,1]$, given by $\kappa(\rho, A) = \rho(A)$, so that $\tilde{P} \preceq M$.

Proposition 1 *Let $M : (X, \mathcal{A}) \to \mathcal{E}(\mathcal{H})$ be a POV measure with commutative range. Let*

$$M(A) = \int_{\mathbb{R}} \nu(x, A)P(dx)$$

be as in Theorem A. Let us define the function $\varphi : \mathbb{R} \to [0,1]^{\mathcal{A}}$ by $\varphi(x) = \nu(x, \cdot)$. Then φ is $B([0,1]^{\mathcal{A}})$-measurable and

$$\tilde{P}(F \cap D) := P(\varphi^{-1}(F)), \quad F \in B([0,1]^{\mathcal{A}})$$

is a PV measure $(D, \mathcal{S}) \to \mathcal{E}(\mathcal{H})$ that satisfies

$$M(A) = \int_D \rho(A)\tilde{P}(d\rho).$$

Proof The function φ is $B([0,1]^{\mathcal{A}})$-measurable, since ν_A is measurable for each A. The rest of the proof is based on the following lemma. □

Lemma 4.1 *Let $F \in B([0,1]^{\mathcal{A}})$ be such that $F \cap D = \emptyset$. Then $P(\varphi^{-1}(F)) = 0$.*

Proof This proof is a modification of the proof of [3, Lemma 2].

So let $f \in [0,1]^{\mathcal{A}}$, then $f \in D$ if and only if $f(A_1) + f(A_2) + \cdots = 1$ for any mutually disjoint sequence $\{A_n\}_n$ of sets in $\mathcal{A}$, such that $\bigcup A_n = X$. For each such sequence, we denote $B(\{A_n\}) := \{f, \sum_n f(A_n) = 1\}$. Clearly, $B(\{A_n\})$ is a closed subset in the compact space $[0,1]^{\mathcal{A}}$ and $D = \cap B(\{A_n\})$, over all such sequences.

Let $F' \subset F$ be closed. Since $F \cap D = \emptyset$, $F' \subset D^c = \bigcup B(\{A_n\})^c$. As F' is compact, there exists a finite number of sequences, $\{A_n^1\}, \ldots, \{A_n^k\}$, such that $F' \subset \bigcup_i B(\{A_n^i\})^c$. From this,

$$\varphi^{-1}(F') \subseteq \bigcup_i \left\{x, \sum_n \nu(x, A_n^i) \neq 1\right\}$$

and the set on the right hand side has $m_0 \circ P$-measure 0, since ν is a weak Markov kernel. It follows that $m_0 \circ P(\varphi^{-1}(F')) = 0$ for all closed sets $F' \subset F$.

Note that the Borel σ-algebra $B([0,1]^{\mathcal{A}})$ is generated by sets of the form $\{f, a_i \leq f(A_i) \leq b_i, i = 1, \dots, m\}$ and these are G_δ-sets. Hence $B([0,1]^{\mathcal{A}})$ is the algebra of Baire sets and consequently, $F \mapsto m_0 \circ P(\varphi^{-1}(F))$ is a Baire measure. Since all Baire measures are regular, we have

$$m_0 \circ P(\varphi^{-1}(F)) = \sup\{m_0 \circ P(\varphi^{-1}(F'),\ F' \subset F, F' \text{ is closed}\} = 0 \qquad \square$$

Proof of Proposition 1 First we prove that $\tilde{P}$ is well defined. Let $F \cap D = E \cap D$, we have to prove that $P(\varphi^{-1}(E)) = P(\varphi^{-1}(F))$. We have $(F \Delta E) = (F \setminus E) \cup (E \setminus F)$, $(F \setminus E) \cap D = F \cap D \setminus E = E \cap D \setminus E = \emptyset$, which entails that $(F \Delta E) \cap D = \emptyset$. Hence $P(\varphi^{-1}(F \Delta E)) = 0$. Since $E = (E \cap F) \cup (E \setminus F)$, $F = (E \cap F) \cup (F \setminus E)$, we obtain $P(\varphi^{-1}(E)) = P(\varphi^{-1}(E \cap F)) = P(\varphi^{-1}(F))$.

For $A \in \mathcal{A}$ then

$$\begin{aligned}
\int_D \rho(A) \tilde{P}(d\rho) &= \int_{[0,1]^{\mathcal{A}}} f(A) P(\varphi^{-1}(df)) \\
&= \int_{\mathbb{R}} \varphi(x)(A) P(dx) \\
&= \int_{\mathbb{R}} \nu(x, A) P(dx) = M(A).
\end{aligned} \qquad \square$$

Proposition 2 *Let M be an observable with commutative range and let $\tilde{P} : (D, \mathcal{S}) \to \mathcal{E}(\mathcal{H})$ be a PV-measure such that*

$$M(A) = \int_D \rho(A) \tilde{P}(d\rho).$$

Then there is a real PV measure $R : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \to \mathcal{E}(\mathcal{H})$ such that $R \sim \tilde{P} \preceq M$. Moreover, $\tilde{P} = \lambda \circ R$ for a weak Markov kernel $\lambda : \mathbb{R} \times \mathcal{S} \to \{0, 1\}$ and $\nu = \kappa \circ \lambda$ satisfies

$$M(A) = \int_{\mathbb{R}} \nu(x, A) R(dx).$$

Proof The range $\mathcal{R}(\tilde{P})$ is a boolean σ-subalgebra of the projection lattice of $\mathcal{H}$. By [10, Theorem 1.6(i)], there is a real PV measure $R : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \to \mathcal{E}(\mathcal{H})$ such that $\mathcal{R}(\tilde{P}) = \mathcal{R}(R)$. By [4, Theorem 3.4], there is a weak Markov kernel $\lambda : \mathbb{R} \times \mathcal{S} \to \{0, 1\}$ such that

$$\tilde{P}(C) = \int_{\mathbb{R}} \lambda(x, C) R(dx), \qquad C \in \mathcal{S},$$

hence $R \preceq \tilde{P}$. As we have seen, $M = \kappa \circ \tilde{P}$ for the Markov kernel $\kappa(\rho, A) = \rho(A)$, $\rho \in D$, $A \in \mathcal{A}$, hence also $\tilde{P} \preceq M$. Now, $M = \nu \circ R$, with $\nu = \kappa \circ \lambda$ according to Sect. 1.

Since $\tilde{P}$ and R have the same ranges, we also have $\tilde{P} \preceq R$ so that $\tilde{P} \sim R$. Notice that by [4, Theorem 3.5(ii′)], there is a function $f : D \to \mathbb{R}$ such that $R(B) = \tilde{P} \circ f^{-1}(B)$, $B \in \mathcal{B}(\mathbb{R})$. □

5 Representation of Commutative POV Measures by Choquet Theorem

Now we consider integral representation of POV measures by Choquet theorem, as it was obtained in [1], but we formulate this result in the setting of Markov kernels. Throughout this section, we suppose that X is a separable metrizable locally compact topological space and $\mathcal{A}$ the σ-algebra of Borel subsets. Note that in this case, X is a Polish space [2, Theorem 4.3.26 and p. 196], so that $(X, \mathcal{A})$ is standard Borel.

We consider the set of all POV measures $(X, \mathcal{A}) \to L_\infty(\mathbb{R}, \tau)$, where τ is a probability measure on $(\mathbb{R}, B(\mathbb{R}))$. By the beginning of the previous section, this set is identified with $\mathcal{K}(\mathbb{R}, \mathcal{A}, \tau)$ and by Theorem 2.1, each of the equivalence classes contains a Markov kernel.

Let $C_0(X)$ be the set of continuous functions $g : X \to \mathbb{C}$, vanishing at infinity and let u_α be an approximate unit in $C_0(X)$. As we have seen in Sect. 2.1, each equivalence class $[\lambda]$ defines a Markov operator

$$T^*_\lambda : C_0(X) \subset L_\infty(X, \mathcal{A}, T_\lambda(\tau)) \to L_\infty(\mathbb{R}, \tau)$$

and, since we may suppose that λ is a Markov kernel,

$$T^*_\lambda(g) = \int_X g(x)\lambda(t, dx), \qquad g \in C_0(X) \tag{6}$$

This implies that $T^*_\lambda(u_\alpha) \to 1$. Conversely, it is known and can be shown for example by the Riesz representation theorem, that any positive linear operator $L : C_0(X) \to L_\infty(\mathbb{R}, \tau)$, satisfying $L(u_\alpha) \to 1$ is given by (6), for some Markov kernel λ.

Let us endow the set $\mathcal{K}(\mathbb{R}, \mathcal{A}, \tau)$ with the topology $\mathcal{T}$, such that $[\lambda_\alpha]$ converges to $[\lambda]$ if

$$\int_{\mathbb{R}} T^*_{\lambda_\alpha}(g)(t)h(t)\tau(dt) \to \int_{\mathbb{R}} T^*_\lambda(g)(t)h(t)\tau(dt)$$

for each $g \in C_0(X)$ and $h \in L_1(\mathbb{R}, \tau)$. Then $\mathcal{K}(\mathbb{R}, \mathcal{A}, \tau)$ is compact and metrizable, this can be proved as [1, Lemma 1], using the fact that the space $L_1(\mathbb{R}, \tau)$ is separable.

5.1 Extreme Points in $\mathcal{K}(\mathbb{R}, \mathcal{A}, \tau)$

Since $\mathcal{K}(\mathbb{R}, \mathcal{A}, \tau)$ is clearly convex, we can apply Choquet theorem (e.g. [6]) to obtain a representing measure, concentrated on the set of extreme points ∂. The extreme points of the set of Markov kernels are the Markov kernels with values in $\{0, 1\}$. Therefore it is not difficult to see that the extreme points in $\mathcal{K}(\mathbb{R}, \mathcal{A}, \tau)$ are the equivalence classes $[\delta]$, with $\delta(t, A) \in \{0, 1\}$, τ-a.s. for all A. We will show that these are precisely the equivalence classes $[\chi_f]$, where $\chi_f(t, A) := \chi_{f^{-1}(A)}(t)$ for a measurable function $f : \mathbb{R} \to X$.

It is clear that any $[\chi_f]$ is an extreme point. The proof of the converse is similar to the proof of [4, Theorem 3.3], so we omit some of the details.

So let $[\delta]$ be an extreme point and let us suppose that δ is a Markov kernel. Let $\pi(A) = \{s, \delta(s, A) = 1\}$, then $\delta(t, A) = \chi_{\pi(A)}(t)$, τ-a.s. Moreover, it can be shown that π is a σ-homomorphism of sets modulo τ. Let $I := \{B \in B(\mathbb{R}),\ \tau(B) = 0\}$, then I is a σ-ideal and $B(\mathbb{R})/I$ is a Boolean σ-algebra. Let $p : B(\mathbb{R}) \to B(\mathbb{R})/I$ be the canonical homomorphism and put $\pi_1 = p \circ \pi$. Then $\pi_1 : \mathcal{A} \to B(\mathbb{R})/I$ is a σ-homomorphism and by [10, Theorem 1.4], there is a measurable function $f : \mathbb{R} \to X$, such that $p \circ \pi(A) = \pi_1(A) = p \circ f^{-1}(A)$. Hence, $\delta(t, A) = \chi_{f^{-1}(A)}(t)$, τ-a.s.

Let now f_1 and f_2 be measurable functions $\mathbb{R} \to X$, such that $[\chi_{f_1}] = [\chi_{f_2}]$. By definition, $\chi_{f_1^{-1}(A)}(t) = \chi_{f_2^{-1}(A)}(t)$, τ-a.s., so that, for each $A \in \mathcal{A}$,

$$\tau(\{s,\ f_1(s) \in A,\ f_2(s) \in A^c\}) = 0$$

Let $B_r(x)$ be the open ball in X with radius r and centered at x, recall that X is metrizable. Then $B_r(x) \in \mathcal{A}$. Since X is separable, we can choose a countable dense subset $\{x_n\}$ in X. It is clear that

$$\{s,\ f_1(s) \neq f_2(s)\} = \bigcup_{n\in\mathbb{N}, r\in\mathbb{Q}} \{s,\ f_1(s) \in B_r(x_n),\ f_2(s) \in B_r(x_n)^c\}$$

so that $f_1 = f_2$, τ-a.s. This implies that the set of extreme points ∂ can be identified with the set $\mathcal{F}(\tau)$ of measurable functions $f : \mathbb{R} \to X$ modulo τ.

Next we identify the topology in $\mathcal{F}(\tau)$, inherited from $\mathcal{T}$. Let $f \in \mathcal{F}(\tau)$, then for any measurable step function $g(x) = \sum_i c_i \chi_{A_i}(x)$, we have

$$T^*_{\chi_f}(g)(t) = \int g(x)\chi_f(t, dx) = \sum_i c_i \chi_{f^{-1}(A_i)}(t) = g \circ f(t)$$

Since $T^*_{\chi_f}$ is normal, this holds for any bounded measurable function g, hence also for all $g \in C_0(X)$.

Let now $f_n, f \in \mathcal{F}(\tau)$ be such that $[\chi_{f_n}] \to [\chi_f]$ in $\mathcal{T}$-topology. Then for all $g \in C_0(X)$, $h \in L_1(\mathbb{R}, \tau)$, we get that

$$\int_{\mathbb{R}} g \circ f_n(t)h(t)\tau(dt) \to \int_{\mathbb{R}} g \circ f(t)h(t)\tau(dt)$$

equivalently,

$$g \circ f_n(t) \to g \circ f(t), \quad \tau\text{-a.s. } g \in C_0(X) \tag{7}$$

It is clear that this holds if $f_n(t) \to f(t)$, τ-a.s. We are going to prove the converse.

So let (7) be true and let s be such that $f_n(s) \nrightarrow f(s)$. Then there is an open neighborhood $f(s) \in O$, such that $f_n(s) \notin O$ for large enough n. Since X is locally compact, there is an open set U, such that $f(s) \in U \subset \bar{U} \subset O$ and $\bar{U}$ is compact. Then $\{f(s)\}$ and U^c are disjoint closed sets.

By Urysohn's Lemma, there is a continuous function $g : X \to [0, 1]$, such that $g(f(s)) = 1$ and $g(x) = 0$ for $x \in U^c$, so that $g(f_n(s)) = 0$ for large enough n, hence $g(f_n(s)) \nrightarrow g(f(s))$. Clearly, $g \in C_0(X)$, since g vanishes outside the compact set $\bar{U}$.

Let us now choose a countable dense subset $\{g_i\}$ in $C_0(X)$ (note that also $C_0(X)$ is separable) and pick a g_i sufficiently close to g, then also $g_i(f_n(s)) \nrightarrow g_i(f(s))$. Therefore,

$$\{s, f_n(s) \nrightarrow f(s)\} = \bigcup_i \{s, g_i \circ f_n(s) \nrightarrow g_i \circ f(s)\}$$

so that $f_n(s) \to f(s)$, τ-a.s. We have shown the following theorem.

Theorem 5.1 *The set ∂ of extreme points of $\mathcal{K}(\mathbb{R}, \mathcal{A}, \tau)$ can be identified with $\mathcal{F}(\tau)$ with the topology of τ-a.s. pointwise convergence.*

5.2 Integral Representations for Commutative POV Measures

Recall that we consider POV measures $(X, \mathcal{A}) \to L_\infty(\mathbb{R}, \tau)$ identified with the set $\mathcal{K}(\mathbb{R}, \mathcal{A}, \tau)$ which is convex, compact and metrizable. The following result is a special case of [1, Theorem 2], taking into account Theorem 5.1.

Theorem 5.2 *For each $\tau \in \mathcal{P}(\mathbb{R}, B(\mathbb{R}))$ and a Markov kernel $\lambda : \mathbb{R} \times A \to [0, 1]$, there is a representing measure σ concentrated on $\mathcal{F}(\tau)$, such that*

$$\lambda(t, A) = \int_{\mathcal{F}(\tau)} \chi_{f^{-1}(A)}(t)\sigma(df), \quad \tau\text{-a.s.}$$

Let now $M : (X, \mathcal{A}) \to B(\mathcal{H})$ be a POV measure with commutative range and let $\mathcal{R} \subset B(\mathcal{H})$ be a commutative subalgebra containing $\mathcal{R}(M)$. As before, we identify $\mathcal{R}$ with $L_\infty(\mathbb{R}, \tau_S)$, $\tau_S = m_0 \circ P_S$, and M with (the $\sim_{\tau_S}$-equivalence class of) a Markov kernel $\lambda : \mathbb{R} \times \mathcal{A} \to [0, 1]$. Let σ be the representing measure of λ on $\mathcal{F}(\tau_S)$. Then we have

$$\begin{aligned} M(A) &= \int_{\mathbb{R}} \lambda(t, A) P_S(dt) = \int_{\mathbb{R}} \int_{\mathcal{F}(\tau_S)} \chi_{f^{-1}(A)}(t)\sigma(df) P_S(dt) \\ &= \int_{\mathcal{F}(\tau_S)} \int_{\mathbb{R}} \chi_{f^{-1}(A)} P_S(dt)\sigma(df) = \int_{\mathcal{F}(\tau_S)} P_S(f^{-1}(A))\sigma(df) \end{aligned}$$

so that M can be obtained as an integral over the set of PV measures on $(X, \mathcal{A})$ with ranges in $\mathcal{R}$.

6 Naimark Extensions

Let $M : (X, \mathcal{A}) \to B(\mathcal{H})$ be a commutative POV measure. Let $\mathcal{K}$ be a Hilbert space containing $\mathcal{H}$ and let $\hat{P}$ be a Naimark extension of M on $\mathcal{K}$, then $E\hat{P}(A)E = M(A)$, where $E : \mathcal{K} \to \mathcal{H}$ is the orthogonal projection.

Let $\mathcal{R}$, S and P_S be as before and let ν be the weak Markov kernel, such that $M(A) = \nu_A(S)$. Let $\hat{R}$ be a real PV measure having the same range as $\hat{P}$, see the proof of Proposition 2, and let $\hat{S}$ be the self-adjoint operator with spectral measure $\hat{R}$.

Then, since $\mathcal{R}(\hat{P}) = \mathcal{R}(\hat{R})$, there is some measurable function $\varphi : X \to \mathbb{R}$, such that $\hat{R}(B) = \hat{P}(\varphi^{-1}(B))$. Then

$$E\hat{R}(B)E = E\hat{P}(\varphi^{-1}(B))E = M(\varphi^{-1}(B))$$

so that $E\chi_B(\hat{S})E = \nu_{\varphi^{-1}(B)}(S)$. Clearly, $(t, B) \mapsto \nu(t, \varphi^{-1}(B))$ is a weak Markov kernel $\mathbb{R} \times B(\mathbb{R}) \to [0, 1]$ with respect to τ_S, which we denote by $\nu \circ \varphi^{-1}$. Then

$$E\chi_B(\hat{S})E = (\nu \circ \varphi^{-1})_B(S) = T^*_{\nu\circ\varphi^{-1}}(\chi_B)(S)$$

Since this holds for any $B \in B(\mathbb{R})$ and $T^*_{\nu\circ\varphi^{-1}}$ is a Markov operator, we get

$$Eg(\hat{S})E = T^*_{\nu\circ\varphi^{-1}}(g)(S)$$

for any bounded measurable function $g : \mathbb{R} \to \mathbb{C}$. Since $(\mathbb{R}, B(\mathbb{R}))$ is standard Borel, there is a Markov kernel $\kappa : \mathbb{R} \times B(\mathbb{R}) \to [0, 1]$ with $T_{\nu\circ\varphi^{-1}} = T_\kappa$, so that

$$Eg(\hat{S})E = T^*_\kappa(g)(S) = \int_{\mathbb{R}} g(t)\kappa(S, dt) = \int_{\mathcal{F}(\tau_S)} g \circ f(S)\sigma(df)$$

where $\kappa = \int_{\mathcal{F}(\tau_S)} \chi_{f^{-1}(\cdot)}\sigma(df)$ is the integral representation from Theorem 5.2.

Let us now suppose that X is a separable metrizable locally compact space and let $M = \lambda \circ P_S$ for a Markov kernel λ. Let $\lambda = \int_{\mathcal{F}(\tau_S)} \chi_{f^{-1}(\cdot)}\sigma(df)$ be the integral representation from Theorem 5.2.

Let $\mathcal{K} = \mathcal{H} \otimes L_2(\mathcal{F}(\tau_S), \sigma)$. Then $\mathcal{K}$ can be identified with the Hilbert space of functions $\varphi : \mathcal{F}(\tau_S) \to \mathcal{H}$, with the inner product $(\varphi, \psi) = \int \langle \varphi(f), \psi(f)\rangle \sigma(df)$. Then $\mathcal{H}$ can be identified with the subspace of constant functions $\mathcal{H} \otimes 1$ and the projection $E : \mathcal{K} \to \mathcal{H} \otimes 1$ is given by

$$(E\varphi)(f) = \int_{\mathcal{F}(\tau_S)} \varphi(g)\sigma(dg), \quad \varphi \in \mathcal{K},\ f \in \mathcal{F}(\tau_S)$$

Let us define a PV measure $\hat{P} : (X, \mathcal{A}) \to B(\mathcal{K})$ by

$$(\hat{P}(A)\varphi)(f) = \chi_{f^{-1}(A)}(S)\varphi(f) = P_S(f^{-1}(A))\varphi(f), \quad \varphi \in \mathcal{K}$$

Then each $\hat{P}(A)$ is a decomposable operator $f \mapsto \hat{P}(A)(f) = P_S(f^{-1}(A))$ and (taking into account the end of previous section) it is rather clear that $\hat{P}$ is a Naimark extension of M.

References

1. Ali, S.T.: A geometrical property of POV-measures and systems of covariance. In: Doebner, H. (ed.) Differential Geometric Methods in Mathematical Physics. Lecture Notes in Math., vol. 905. Springer, Berlin (1982)
2. Engelking, R.: General Topology. Heldermann, Berlin (1989)
3. Holevo, A.S.: An analogue of the theory of statistical decisions in noncommutative probability theory. Trans. Mosc. Math. Soc. **26**, 133–147 (1972)

4. Jenčová, A., Pulmannová, S.: How sharp are PV measures? Rep. Math. Phys. **59**, 257–266 (2007)
5. Jenčová, A., Pulmannová, S., Vinceková, E.: Sharp and Fuzzy observables on effect algebras. Int. J. Theor. Phys. **47**, 125–148 (2008)
6. Phelps, R.R.: Lectures on Choquet's Theorems. Van Nostrand, Princeton/New Jersey (1966)
7. Riesz, F., Nagy, B.Sz.: Leçons d'Analyse Fonctionelle. Academie des Sciences de Hongrie, Szeged (1955)
8. Štěpán, J.: Probability Theory. Academia, Praha (1987). (Teorie Pravděpodobnsti, in Czech)
9. Strasser, H.: Mathematical Theory of Statistics. de Gruyter, Berlin (1985)
10. Varadarajan, V.S.: Geometry of Quantum Theory. Springer, New York (1985)
11. von Neumann, J.: Zur Algebra der Functionaloperatoren und Theorie der normalen Operatoren. Math. Ann. **102**, 370–427 (1929)