TCM315 Fall 2022: Introduction to Open Quantum Systems

Lecture 6: Elements of Quantum Probability

Course handouts are designed as a study aid and are not meant to replace the recommended textbooks. Handouts may contain typos and/or errors. The students are encouraged to verify the information contained within and to report any issue to the lecturer.

CONTENTS

INTRODUCTION

By Quantum Probability we mean how Quantum Theory can be used to define probabilistic models satisfying the Kolmogorov' s axioms [6]. An introduction for physicists is chapter 2 of [2]. An elementary discussion of the geometry of the space of state operators is in chapter 3 of [9].

We discuss the celebrated "no hidden variable" theorem by von Neumann [11] later generalized on the basis of weaker assumptions for Hilbert spaces of more than two dimensions by Gleason [5] and by Busch [4]. The physical relevance of the "no hidden variable" theorem has been discussed by Bell in "On the problem of hidden variables in quantum mechanics" article 1 in [1] suggesting that proof fails to exclude any significant class of hidden variables. An interesting re-assessment of von Neumann theorem and Bell's criticism is contained in the article by J. Bub [3] to which these notes are also beholden.

Unless otherwise explicitly stated, we always work in a finite dimensional Hilbert space $\mathcal{H}$ with dimension

$$d = \dim \mathcal{H}$$

A BRIEF REMINDER OF PROBABILITY THEORY

An elementary probabilistic model is a conceptualization of an **experiment**. The result of the experiment is to produce exactly one outcome out of a set of possible **elementary events** or **atoms**. The collection of all possible outcomes is called the **sample space** which we call Ω. Probability is then encoded in the following Kolmogorov's axioms [6, 7]:

A-i The collection $\mathcal{F}$ of all the subsets of Ω forms an algebra of sets. Such algebra includes Ω, $\emptyset$ and all sets, called **events**, that can be formed by composing elementary events.

A-ii To any event $A \in \mathcal{F}$ it is possible to associate a **non negative** real number $\mathrm{P}(A)$, the **probability** of the event A.

A-iii $\mathrm{P}(\Omega) = 1$

A-iv If A and B have no element in common ($A \cap B = \emptyset$ implies $A \cup B \equiv A + B$)

$$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$$

If $A \subseteq \Omega$ we denote by $\bar{A}$ its **complement**

$$A + \bar{A} = \Omega$$

A-iii and **A-iv** imply that

$$\mathrm{P}(A) + \mathrm{P}(\bar{A}) = 1$$

In particular, since $\bar{\Omega} = \emptyset$ and $\mathrm{P}(\bar{A}) \geq 0$ immediate consequences are

$$\mathrm{P}(\emptyset) = 0\,, \qquad 0 \leq \mathrm{P}(A) \leq 1$$

The axioms above must be strengthened if we require $\mathcal{F}$ to be closed under *countable* union. We talk in such a case of σ-algebra and we need to replace **A-iv** with

A-v The probability of the *infinite* union of disjoint events equals the sum of their probabilities

$$\mathrm{P}\left(\sum_i^{\infty} F_i\right) = \sum_{i=1}^{\infty} \mathrm{P}(F_i)$$

Conditional Probability

An important concept is that of **conditional probability**. If $\mathrm{P}(A) > 0$ we can write for any $B \in \mathcal{F}$

$$\mathrm{P}(B|A) = \frac{\mathrm{P}(B \cap A)}{\mathrm{P}(A)}$$

In order to appreciate the meaning of this definition we can reason as follows. Let N be the number of the possible results of an experiment and let N_1 N_2 and N_{12} respectively the number of times that the events F_1, F_2 and $F_1 \cap F_2$ are observed. Then, if we identify the probabilities with frequencies i.e.

$$\mathrm{P}(F_1) = \frac{N_1}{N}$$

then

$$\mathrm{P}(F_1|F_2) = \frac{N_{12}}{N}\frac{N}{N_2} = \frac{N_{12}}{N_2}$$

which is the frequency of $F_1 \cap F_2$ if we restrict the attention to the number of results for which the event F_2 occurs. An important consequence of the definition of conditional probability is **Bayes's formula**:

$$\mathrm{P}(A|B) = \mathrm{P}(B|A)\frac{\mathrm{P}(A)}{\mathrm{P}(B)}$$

BRIDGE RELATION BETWEEN QUANTUM MECHANICS AND PROBABILITY THEORY

The postulates of Quantum Mechanics surmise the following bridge formula between quantum theory and classical probability. The expectation value of a dynamical variable $\mathscr{A}$ must be computed as the trace of the product of the self-adjoint operator $\mathbb{A}$ associated to $\mathscr{A}$ and the state operator $\boldsymbol{\rho}$ specifying the probability ensemble describing the physical model:

$$\mathrm{E}\,\mathscr{A} = \mathrm{Tr}\,(\boldsymbol{\rho}\mathbb{A}) \tag{1}$$

Thus assigning $\boldsymbol{\rho}$ specifies the probability measure which we need to use to weigh the occurrence of events. Hence, we have

$$1 = \mathrm{P}(\Omega\big|\boldsymbol{\rho}) \equiv \mathrm{Tr}\,\boldsymbol{\rho}$$

whereas measurable events corresponds to eigenvalues of self-adjoint operators:

$$\mathscr{A} \qquad \Leftrightarrow \qquad \mathbb{A} = \sum_{i=1}^{\mathscr{N}} a_i \mathbb{P}_{a_i}$$

yields

$$\mathrm{P}\left(\mathscr{A} = \alpha_i\big|\boldsymbol{\rho}\right) = \mathrm{Tr}(\boldsymbol{\rho}\,\mathbb{P}_{a_i})$$

Conditional Probability

If ρ is a mixture,

$$\boldsymbol{\rho} = \sum_{k=1}^{\mathscr{M}} \wp_k\, \boldsymbol{\rho}_k$$

We may think of the collection $\{\boldsymbol{\rho}_k\}_{k=1}^{\mathscr{M}}$ as the state operators obtained by performing the generalized measurement specified by the operators $\{\mathbb{M}_k\}_{k=1}^{\mathscr{M}}$ on an initial ensemble $\boldsymbol{\rho}$

$$\boldsymbol{\rho}_k = \frac{\mathbb{M}_k\boldsymbol{\rho}\mathbb{M}_k^\dagger}{\wp_k} \qquad \& \qquad \wp_k = \mathrm{Tr}\left(\mathbb{M}_k\boldsymbol{\rho}\mathbb{M}_k^\dagger\right)$$

Thus

$$\mathrm{P}(\mathscr{A} = a_i|\boldsymbol{\rho}_k, \boldsymbol{\rho}) = \mathrm{Tr}(\boldsymbol{\rho}_k\,\mathbb{P}_{a_i})$$

The self-consistence of the definition is granted because

$$\sum_{i=1}^{\mathscr{N}} \mathrm{P}(\mathscr{A} = a_i|\boldsymbol{\rho}_k, \boldsymbol{\rho}) = \mathrm{Tr}(\boldsymbol{\rho}_k) = 1$$

Bayes formula

Once we recovered the notion of conditional probability in the quantum context, we can derive Bayes formula

$$\mathrm{P}(\boldsymbol{\rho}_k|\mathscr{A} = a_i, \boldsymbol{\rho}) = \frac{\mathrm{P}(\mathscr{A} = a_i|\boldsymbol{\rho}_k, \boldsymbol{\rho})\,\wp_k}{\mathrm{P}(\mathscr{A} = a_i|\boldsymbol{\rho})} = \frac{\mathrm{Tr}(\boldsymbol{\rho}_k\,\mathbb{P}_{a_i})\wp_k}{\mathrm{Tr}(\boldsymbol{\rho}\,\mathbb{P}_{a_i})}$$

which is again well defined since

$$\sum_{k=1}^{\mathscr{M}} \mathrm{P}(\boldsymbol{\rho}_k|\mathscr{A} = a_i, \boldsymbol{\rho}) = \sum_{k=1}^{\mathscr{M}} \frac{\mathrm{Tr}(\boldsymbol{\rho}_k\,\mathbb{P}_{a_i})\wp_k}{\mathrm{Tr}(\boldsymbol{\rho}\,\mathbb{P}_{a_i})} = 1$$

PROPERTIES OF ENSEMBLES IN QUANTUM PROBABILITY

We now turn to discuss some properties of characteristic of probability ensembles derived in accordance with the postulates of quantum mechanics. In order to emphasize their peculiarity we start by contrasting the corresponding notions in probability ensembles specified by probability density functions.

Absence of dispersion free ensembles

We contast the situation with the classical case. To this goal we resort to the notion of **random variable**. By random variable we mean any measurable function mapping the sample space into real numbers.

Classical case

It easy to describe deterministic i.e. non-random or "dispersion free" quantities in classical probability theory. A random variable ξ with probability

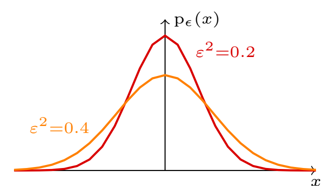
$$\Pr(x \, \leq \, \xi < x + \mathrm{d}x) = \mathrm{p}_{a,\varepsilon}(x)\mathrm{d}x$$

is *dispersion free* if it takes with probability one the value, say, $\xi = a$. We can also associate to ξ a probability measure by considering the sequence of probability densities

$$\mathrm{p}_{a,\varepsilon}(x) = \frac{e^{-\frac{(x-a)^2}{2\,\varepsilon^2}}}{\sqrt{2\,\pi\,\varepsilon^2}}$$

and correspondingly the one parameter ε family of random variables such that

$$\Pr(\xi \in I) = \int_I \mathrm{d}x \, \frac{e^{-\frac{(x-a)^2}{2\,\varepsilon^2}}}{\sqrt{2\,\pi\,\varepsilon^2}}$$



Gaussian for two values of ε showing the "bell shaped" curve shrinking and peaking around $a = 0$ as ε decreases

Then for any bounded function $f\colon \mathbb{R} \mapsto \mathrm{R}$

$$\mathrm{E}\, f(\xi_\varepsilon) = \int_{\mathbb{R}} \mathrm{d}x \, \frac{e^{-\frac{(x-a)^2}{2\,\varepsilon^2}}}{\sqrt{2\,\pi\,\varepsilon^2}} \, f(x) = \int_{\mathbb{R}} \mathrm{d}x \, \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\,\pi}} \, f(a + \varepsilon\, x)$$

As the integral is bounded for any ε we can carry the limit over the integral sign and conclude

$$\mathrm{E}\, f(\xi) = \lim_{\varepsilon \downarrow 0} \mathrm{E}\, f(\xi_\varepsilon) = f(a)$$

or equivalently

$$\mathrm{E}\, f(\xi) = f\,(\mathrm{E}\,\xi)$$

independently of the choice of a bounded f. Interestingly, this is not possible in Quantum Mechanics.

Quantum case

Proposition. *There is no state operator $\boldsymbol{\rho}$ such that for **any** function f and **any** self-adjoint operator $\mathbb{A}$, the chain of identities*

$$\mathrm{E}\, f(\mathscr{A}) = \mathrm{Tr}(\boldsymbol{\rho}\, f(\mathbb{A})) = f(\mathrm{Tr}(\boldsymbol{\rho}\,\mathbb{A})) = f(\mathrm{E}\,\mathscr{A}) \tag{2}$$

holds true.

Proof.
The existence of a dispersion free state, entails that (2) must hold true for any $\mathbb{A}$ and any f. The proof of the claim proceeds by showing that assuming (2) leads to a contradiction.

We preliminarily observe that even if $\boldsymbol{\rho}$ describes a pure state, in general the pure state won't be an eigenstate of a generically chosen self-adjoint operator $\mathbb{A}$. In fact, we can take $\mathbb{A}$ to be the projector on an arbitrary unit vector $\boldsymbol{a}$

$$\mathbb{A} = \boldsymbol{a}\boldsymbol{a}^{\dagger}$$

Furthermore, we can choose $f(x) = x^2$. Then if (2) holds we must have

$$\mathrm{Tr}(\boldsymbol{\rho}\,\mathbb{A}^2) = (\mathrm{Tr}(\rho\,\mathbb{A}))^2$$

Since by hypothesis $\mathbb{A}$ is a projector

$$\mathrm{Tr}(\boldsymbol{\rho}\,\mathbb{A}^2) = \mathrm{Tr}(\boldsymbol{\rho}\,\mathbb{A})$$

we arrive to the equation

$$\mathrm{Tr}(\boldsymbol{\rho}\,\mathbb{A}) = (\mathrm{Tr}(\boldsymbol{\rho}\,\mathbb{A}))^2$$

The equation admits solution only if

$$\boldsymbol{\rho} = 0 \qquad \text{or} \qquad \boldsymbol{\rho} = \mathbb{1}_d$$

Both solutions contradict the hypothesis that $\boldsymbol{\rho}$ is a state matrix since they do not satisfy $\mathrm{Tr}\,\boldsymbol{\rho} = 1$ □

The absence of dispersion free ensembles was from the inception of Quantum Mechanics a feature which was challenging the interpretation. Many physicists adhere(d) to the **Principle of Sufficient Reason** (see e.g. [8]) often reported as stating that

For every fact F, there must be a sufficient reason why F is the case.

In an experimental setting this means that measurement performed in exactly the same conditions should produce the same results. If they do not, then the principle implies the existence of one or more non-controlled "hidden variable" responsible for the plurality of registered outcomes. The absence of dispersion-less ensemble implied by (1) hints at the violation of the principle. The proof by von Neumann in [11] (see section below) that the bridge relation is the only possible without substantially modifying the postulates, exhibits the unavoidably random character of experiment outcomes entailed by the assumption that all what we can say about dynamical variables must be encoded into self-adjoint operators acting on an Hilbert space.

Heisenberg's uncertainty relations

We recall that the variance of a random variable ξ is

$$\mathrm{Var}\,\xi = \mathrm{E}\,(\xi - \mathrm{E}\,\xi)^2 = \mathrm{E}\,\xi^2 - (\mathrm{E}\,\xi)^2 \,\geq\, 0$$

By definition the variance vanishes if and only if the the ensemble is dispersion-less The variance of a dynamical variable embodied in Quantum Mechanics by a self-adjoint operator $\mathbb{X}$ is then

$$\mathrm{Var}\,\xi = \mathrm{Tr}\,\boldsymbol{\rho}\,(\mathbb{X} - (\mathrm{Tr}\,(\boldsymbol{\rho}\mathbb{X}))\,\mathbb{1}_d)^2$$

Proposition. *Let $\mathscr{A}_1$ and $\mathscr{A}_2$ two dynamical variables respectively represented by the self-adjoint operators $\mathbb{A}_1$ and $\mathbb{A}_2$. Then*

$$\sqrt{\left(\mathrm{Var}\,\mathscr{A}_1\right)\left(\mathrm{Var}\,\mathscr{A}_2\right)} \,\geq\, \frac{1}{2}\,\mathrm{Tr}\,(\boldsymbol{\rho}\,[\mathbb{A}_1\,,\mathbb{A}_2]) \tag{3}$$

Proof. Let us define

$$\tilde{\mathbb{A}}_i = \mathbb{A}_i - (\operatorname{Tr}(\boldsymbol{\rho}\mathbb{A}_i))\,\mathbb{1}_d \qquad i = 1.2$$

Then, by definition

$$\big(\operatorname{Var} A_1\big)\big(\operatorname{Var} A_2\big) \equiv \big(\operatorname{E}\tilde{A}_1^2\big)\big(\operatorname{E}\tilde{A}_2^2\big) = \big(\operatorname{Tr}\boldsymbol{\rho}\tilde{\mathbb{A}}_1^2\big)\big(\operatorname{Tr}\boldsymbol{\rho}\tilde{\mathbb{A}}_2^2\big)$$

We notice that for any $\mathbb{X}_1, \mathbb{X}_2 \in \mathcal{M}_d(\mathbb{C})$ the quantity

$$\langle\, \mathbb{X}_1\,, \mathbb{X}_2\,\rangle_{\boldsymbol{\rho}} \equiv \operatorname{Tr}\left(\boldsymbol{\rho}\mathbb{X}_1^\dagger\,\mathbb{X}_2\right)$$

defines an inner product. Schwartz's inequality then yields

$$\big(\operatorname{Tr}\boldsymbol{\rho}\tilde{\mathbb{A}}_1^2\big)\big(\operatorname{Tr}\boldsymbol{\rho}\tilde{\mathbb{A}}_2^2\big) \geq \left|\operatorname{Tr}\boldsymbol{\rho}\,\tilde{\mathbb{A}}_1\tilde{\mathbb{A}}_2\right|^2$$

Since by hypothesis $\mathbb{A}_1, \mathbb{A}_2$ are **self-adjoint**

$$\tilde{\mathbb{A}}_1\,\tilde{\mathbb{A}}_2 = \frac{\tilde{\mathbb{A}}_1\,\tilde{\mathbb{A}}_2 + \tilde{\mathbb{A}}_2\,\tilde{\mathbb{A}}_1}{2} + \frac{\tilde{\mathbb{A}}_1\,\tilde{\mathbb{A}}_2 - \tilde{\mathbb{A}}_2\,\tilde{\mathbb{A}}_1}{2} = \operatorname{Re}\left(\tilde{\mathbb{A}}_1\tilde{\mathbb{A}}_2\right) + \imath\operatorname{Im}\left(\tilde{\mathbb{A}}_1\tilde{\mathbb{A}}_2\right)$$

As a consequence, we get into the chain of relations

$$\begin{aligned}\left|\operatorname{Tr}\boldsymbol{\rho}\,\tilde{\mathbb{A}}_1\tilde{\mathbb{A}}_2\right|^2 &= \left(\operatorname{Re}\operatorname{Tr}\boldsymbol{\rho}\,\tilde{\mathbb{A}}_1\tilde{\mathbb{A}}_2\right)^2 + \left(\operatorname{Im}\operatorname{Tr}\boldsymbol{\rho}\,\tilde{\mathbb{A}}_1\tilde{\mathbb{A}}_2\right)^2 \\ &\geq \left(\operatorname{Im}\operatorname{Tr}\boldsymbol{\rho}\,\tilde{\mathbb{A}}_1\tilde{\mathbb{A}}_2\right)^2 = \frac{1}{4}\left(\operatorname{Tr}\boldsymbol{\rho}\,\left[\tilde{\mathbb{A}}_1\,,\tilde{\mathbb{A}}_2\right]\right)^2 = \frac{1}{4}\left(\operatorname{Tr}\boldsymbol{\rho}\,[\mathbb{A}_1\,,\mathbb{A}_2]\right)^2\end{aligned}$$

The last equality stems from the fact that the identity matrix commutes with any other matrix. We therefore proved the inequality

$$\big(\operatorname{Var}\mathscr{A}_1\big)\big(\operatorname{Var}\mathscr{A}_2\big) \geq \frac{1}{4}\left(\operatorname{Tr}\boldsymbol{\rho}\,[\mathbb{A}_1\,,\mathbb{A}_2]\right)^2$$

whence the claim readily follows. □

The physical interpretation of the uncertainty principle is that if we prepare a large number of quantum systems in the state described by the same state operator $\boldsymbol{\rho}$, and then perform measurements of the dynamical variable $\mathscr{A}_1$ on some of those systems, and of $\mathscr{A}_2$ in others, then the product of the variances will satisfy the uncertainty relation (3). The lower bound on the uncertainty relation vanishes if the operators commute. A remarkable result in spectral theory (appendix) tells us that self-adjoint operators commute **if and only if** are simultaneously diagonalizable. From the physics slant, this means that a projective measurement onto an eigenspace of say $\mathbb{A}_1$, followed by a projective measurement onto an eigenspace of $\mathbb{A}_2$ provides the same result as if the order of the measurement is inverted. We may rephrase this by saying that $\mathscr{A}_1$ and $\mathscr{A}_2$ are simultaneously measurable.

VON NEUMANN'S "NO HIDDEN VARIABLES" THEOREM

In the inquire of the relation between the postulates of Quantum Mechanics and probabilistic models, a general question may arise whether it might be possible to associate to the postulates a different rule to compute the probability of measurement outcomes. von Neumann already addressed this point already in 1932 (chapter IV of [11]). von Neumann's answer is negative and goes under the name of "no-hidden variable theorem" von Neumann starts from general considerations

> *Let us forget the whole of quantum mechanics but retain the following. Suppose a system* S *is given, which is characterized for the experimenter by the enumeration of all the effectively measurable quantities in it and their functional relations with one another. With each quantity we include the directions as to how it is to be measured – and how its value is to be read or calculated from the indicator positions on the measuring instruments. If $\mathscr{A}$ is a quantity and $f(x)$ any function, then the quantity $f(\mathscr{A})$ is defined as follows: To measure $f(\mathscr{A})$, we measure $\mathscr{A}$ and find the value a (for $\mathscr{A}$). Then $f(\mathscr{A})$ has the value $f(a)$. As we see, all quantities $f(\mathscr{A})$ ($\mathscr{A}$ fixed, $f(x)$ an arbitrary function) are measured simultaneously with $\mathscr{A}$ .... But it should be realized that it is completely meaningless to try to form $f(\mathscr{A}, \mathscr{B})$ if $\mathscr{A}, \mathscr{B}$ are not simultaneously measurable: there is no way of giving the corresponding measuring arrangement.* [11] (pag. 297 of the English translation).

and then proceeds to formulate working hypotheses under which he derives his "no-go" result.

Hypotheses about dynamical variables

A If the dynamical variable $\mathscr{A}$ is by nature non-negative, then

$$\mathrm{E}\,\mathscr{A} \geq 0$$

and in particular

$$\mathrm{E}(1) = 1$$

B if $\{\mathscr{A}_i\}_{i\geq 1}$ are arbitrary dynamical variables and $\{c_i\}_{i\geq 1}$ are real numbers, then

$$\mathrm{E}\left(\sum_{i=1} c_i\,\mathscr{A}_i\right) = \sum_{i=1} c_i\;\mathrm{E}\,(\mathscr{A}_i)$$

von Neumann emphasizes ([11] pag.309 of the English translation) that if the quantities $\{\mathscr{A}_i\}_{i\geq 1}$ are simultaneously measurable, then $\sum_{i=1} c_i\,\mathscr{A}_i$ is the ordinary sum, but if they are not simultaneously measurable, then "*the sum is characterized ... only in an implicit way*". The **hypothesis that the sum may be given a meaning also for non commuting quantities** corresponds to the to the intuitive idea that the expectation value of a kinetic plus potential Hamiltonian may be computed as the average of the kinetic plus the average of the potential. From the mathematical point of view, is a **strong requirement** that later development managed to considerably weaken (see section below).

Hypotheses about correspondence between dynamical variables and operators

Two further assumptions relate physical quantities to Hilbert space operators.

I If the dynamical variable $\mathscr{A}$ has the operator $\mathbb{A}$, then the quantity $f(\mathscr{A})$ has the operator $f(\mathbb{A})$.

II if the dynamical variables $\{\mathscr{A}_i\}_{i\geq 1}$ have the operators $\{\mathbb{A}_i\}_{i\geq 1}$ then $\sum_{i\geq 1}\mathscr{A}_i$ has the operator $\sum_{i\geq 1}\mathbb{A}_i$.

Statement and proof of the theorem

von Neumann's "no-hidden variable" theorem consists in proving that **if** (**A**), (**B**), (**I**), and (**II**) hold true **then the expectation value function is uniquely defined by the trace function**:

$$\mathrm{E}\,\mathscr{A} = \mathrm{Tr}(\boldsymbol{\rho}\,\mathbb{A})$$

where $\boldsymbol{\rho}$ is a self-adjoint operator independent of $\mathbb{A}$ enjoying the properties

- $\boldsymbol{\rho} = \boldsymbol{\rho}^{\dagger}$ (self-adjoint).
- $\mathrm{Sp}\,\boldsymbol{\rho} \geq 0$ (positive definite).
- $\mathrm{Tr}\,\boldsymbol{\rho} = 1$ (unital).

The proof goes as follows.

Proof.
First we select a self-adjoint basis $\{\mathbb{H}_i\}_{i=1}^{d^2}$ of $\mathcal{H}$ orthonormal with respect to the Hilbert-Schmidt inner product

$$\mathbb{H}_i^{\dagger} = \mathbb{H}_i \tag{4a}$$

$$\mathrm{Tr}\,\mathbb{H}_i\mathbb{H}_j = \delta_{ij} \tag{4b}$$

In such a basis any self-adjoint operator $\mathbb{A}$ takes the form

$$\mathbb{A} = \sum_{i=1}^{d^2} a_i \mathbb{H}_i$$

where $a_i \in \mathbb{R}$ for any $\mathbb{R}$. According to (**II**) we surmise that

$$\mathscr{A} = \sum_{i=1}^{d^2} a_i \mathscr{H}_i$$

We then use (**II**) and (**B**) to arrive at

$$\mathrm{E}\,\mathscr{A} = \sum_{i=1}^{d^2} a_i\,\mathrm{E}\,\mathscr{H}_i$$

As the $\{\mathrm{E}\,\mathscr{H}_i\}_{i=1}^{d^2}$ are real numbers, we can readily find a self-adjoint matrix by setting

$$\boldsymbol{\rho} = \sum_{i=1}^{d^2} r_i\,\mathbb{H}_i \qquad \& \qquad r_i = \mathrm{E}\,\mathscr{H}_i$$

such that the identity

$$\mathrm{E}\,\mathscr{A} = \sum_{i=1}^{d^2} r_i\,a_i = \mathrm{Tr}(\boldsymbol{\rho}\mathbb{A})$$

holds true in consequence of (4). The arbitrariness in the choice of the basis $\{\mathbb{H}_i\}_{i=1}^{d^2}$ implies that the coefficients $\{r_i\}_{i=1}^{d^2}$ are independent of the coefficients $\{a_i\}_{i=1}^{d^2}$ and hence $\boldsymbol{\rho}$ of $\mathbb{A}$. It now remains to derive which sort of restrictions (**A**) imposes on $\boldsymbol{\rho}$.

- $\boldsymbol{\rho}$ is **positive definite**. Let $\mathscr{A}$ be associated to the projectors upon the state ψ:

  $$\mathbb{A} = \boldsymbol{\psi}\boldsymbol{\psi}^\dagger$$

  Clearly for any $\boldsymbol{v}$ in the Hilbert space $\mathcal{H}$

  $$\boldsymbol{v}^\dagger \mathbb{A} \boldsymbol{v} = \left\|\boldsymbol{v}^\dagger \boldsymbol{\psi}\right\|^2 \geq 0$$

  This means that $\mathscr{A}$ is positive definite. By (**A**) we must have

  $$0 \leq \mathrm{E}\,\mathscr{A} = \mathrm{Tr}\,\boldsymbol{\rho}\mathbb{A} = \boldsymbol{\psi}^\dagger \boldsymbol{\rho} \boldsymbol{\psi}$$

  The inequality must hold for any $\mathbb{A}$ of projector form hence we conclude that $\boldsymbol{\rho}$ must be positive definite.

- $\boldsymbol{\rho}$ is **unital**: the hypothesis (**A**) immediately implies

  $$1 = \mathrm{E}(1) = \mathrm{Tr}(\boldsymbol{\rho}\,\mathbb{1}_d)$$

□

Drawing from [3] we emphasize the following implications of the theorem

- The assumptions (**A**), (**B**), (**I**), and (**II**) uniquely determine the **bridge relation** (1) between quantum mechanics and probability theory.
- The bridge relations bring as immediate consequences the absence of dispersion free ensembles and uncertainty relations.
- Hence it is impossible to construct an hidden variable theory able to overcome the random nature of measurement outcomes by associating the same deterministic value to a dynamical variable measured under identical measurement conditions while **at the same time** maintaining the correspondence between observables and the self-adjoint operators surmised in the postulates of quantum mechanics.

Developments

Gleason's extension of von Neumann's theorem eminently consists in relaxing the hypothesis (II) by requiring it to hold only for **simultaneously measurable quantities**. To explain this point we emphasize that projectors onto orthogonal sub-spaces trivially commute. Gleason introduced the notion of **frame function**

Definition. *(**Gleason's frame function**) as a **function over the set of projectors** over an Hilbert space $\mathcal{H}$ such that*

i $1 \geq f_G(\boldsymbol{v}\boldsymbol{v}^\dagger) \geq 0$ *for any unit vector* $\boldsymbol{v} \in \mathcal{H}$.

ii $f_G(\mathbb{1}_d) = 1$

iii if $\{\boldsymbol{v}_i\}_{i=1}^d$ *is an **orthonormal basis** of* $\mathcal{H}$ *then*

$$\sum_{i=1}^{d} f_G(\boldsymbol{v}_i\boldsymbol{v}_i^\dagger) = f_G\left(\sum_{i=1}^{d} \boldsymbol{v}_i\boldsymbol{v}_i^\dagger\right) \tag{5}$$

The key point is that additivity is required only on orthonormal projectors which are thus **commuting operators**. Physically, this means having set the focus on how to attribute probabilities to projective measurements.

Theorem (Gleason's)**.** *Let be* f_G *a frame function **over the set of projectors** acting on a real or complex Hilbert space* $\mathcal{H}$ *of dimension* $d > 2$*, Then there exists a state operator* $\boldsymbol{\rho}$ *such that*

$$f_G(\boldsymbol{v}\boldsymbol{v}^\dagger) = \mathrm{Tr}\left(\boldsymbol{\rho}\,\boldsymbol{v}\boldsymbol{v}^\dagger\right) = \boldsymbol{v}^\dagger\boldsymbol{\rho}\,\boldsymbol{v}$$

*for **all unit vector*** $\boldsymbol{v} \in \mathcal{H}$.

The theorem dates to 1957 [5]. The proof is notoriously complicated. A thorough discussion of the theorem and its implications as well as a simpler proof holding for $d = 3$ dimensional Hilbert spaces can be found in [10]. Later Busch [4] showed that the proof can be greatly simplified if we consider frame functions defined on the **effects** of an arbitrary generalized measurement rather than on projectors spanning the Hilbert space. The starting point is an alternative definition of frame function.

Definition. *(**Busch's frame function**) A frame function is a function* f_B *mapping an effect acting on a Hilbert space* $\mathcal{H}$ *to* $\mathbb{R}_+$ *and satisfying the following properties*

i $1 \geq f_B(\mathbb{E}) \geq 0$ *for any effect* $\mathbb{E}$

ii $f_B(\mathbb{1}_d) = 1$

iii $f_B\left(\sum_{i=1}^n \mathbb{E}_i\right) = \sum_{i=1}^n f_B(\mathbb{E}_i)$ *for any collection of effects such that* $\mathrm{Sp}\left(\sum_{i=1}^n \mathbb{E}_i\right) \leq 1$

Busch's definition is aimed at attributing a probability measure to generalized and not just to projective measurements.

Theorem (**Busch's**)**.** *Any generalized probability measure specified by a frame function is of the form*

$$f_B(\mathbb{E}) = \mathrm{Tr}\left(\boldsymbol{\rho}\,\mathbb{E}\right)$$

for some state operator $\boldsymbol{\rho}$.

APPENDIX

Commuting self-adjoint operators

The uncertainty relation becomes trivial if the operators commutes. From the mathematics slant this condition has an important consequence

Theorem. *Let $\mathbb{A}_1, \mathbb{A}_2 \in \mathcal{M}_d(\mathbb{C})$ be self-adjoint. Then they are simultaneously diagonalizable if and only if they commute.*

Proof.

- **Necessary condition**. Suppose $\mathbb{A}_1$, $\mathbb{A}_2$ are self-adjoint and simultaneously diagonalizable. This means that there exists an unitary matrix U such that

$$\mathbb{A}_i = \mathbb{U}\,\mathrm{Diag}(\mathbb{A}_i)\,\mathbb{U}^\dagger \qquad i = 1, 2$$

Diag($\mathbb{A}_i$) stands for the diagonal form of $\mathbb{A}_i$. Then

$$\begin{aligned}[\mathbb{A}_1, \mathbb{A}_2] &= \mathbb{U}\,\mathrm{Diag}(\mathbb{A}_1)\,\mathbb{U}^\dagger\,\mathbb{U}\,\mathrm{Diag}(\mathbb{A}_2)\,\mathbb{U}^\dagger - \mathbb{U}\,\mathrm{Diag}(\mathbb{A}_2)\,\mathbb{U}^\dagger\,\mathbb{U}\,\mathrm{Diag}(\mathbb{A}_1)\,\mathbb{U}^\dagger \\ &= \mathbb{U}\,[\mathrm{Diag}(\mathbb{A}_1)\,,\,\mathrm{Diag}(\mathbb{A}_2)]\,\mathbb{U}^\dagger = 0\end{aligned}$$

- **sufficient condition**. Suppose $\mathbb{A}_1$, $\mathbb{A}_2$ are self-adjoint and commute. Let

$$\mathbb{A}_1\,\boldsymbol{a}_{1,k} = a_{1,k}\,\boldsymbol{a}_{1,k}$$

then

$$\mathbb{A}_1\,\mathbb{A}_2\,\boldsymbol{a}_{1,k} = \mathbb{A}_2\,\mathbb{A}_1\,\boldsymbol{a}_{1,k} = a_{1,k}\,\mathbb{A}_2\,\boldsymbol{a}_{1,k}$$

If a_1 is non degenerate it follows that

$$\mathbb{A}_2\,\boldsymbol{a}_{1,k} = c\,\boldsymbol{a}_{1,k}$$

must hold for some $c \in \mathbb{C}$. But in such a case c is an eigenvalue of $\mathbb{A}_2$ and $\mathbb{A}_2$ is self-adjoint then $c = a_{2,k} \in \mathbb{R}$. If the eigenvalues $a_{1,k_1} = \cdots = a_{1,k_j}$ for $j > 1$ are degenerate, the theorem's claim follows because

$$\tilde{\mathbb{B}} = \boldsymbol{a}^\dagger_{1,k_i}\mathbb{A}_2\boldsymbol{a}_{1,k_j}$$

forms a self-adjoint $j \times j$ matrix. Such matrix is also diagonalizable by an orthonormal linear combination of the eigenvectors $\boldsymbol{a}_{1,k_1}\ \ldots, \boldsymbol{a}_{1,k_j}$ of $\mathbb{A}_1$.

□

* * *

[1] J. S. Bell. *Speakable and unspeakable in quantum mechanics: collected papers on quantum philosophy.* Scence. Cambridge University Press, 2, revised edition, 2004.

[2] H.-P. Breuer and F. Petruccione. *The Theory of Open Quantum Systems.* Oxford University Press, reprint edition, 2002.

[3] J. Bub. Von Neumann' s "No Hidden Variables" Proof: A Re-Appraisal. *Foundations of Physics*, 40(9–10):1333–1340, June 2010.

[4] P. Busch. Quantum States and Generalized Observables: A Simple Proof of Gleason's Theorem. *Physical Review Letters*, 91(12):120403, September 2003.

[5] A. Gleason. Measures on the Closed Subspaces of a Hilbert Space. *Journal of Mathematics and Mechanics*, 6(6):885–893, 1957.

[6] A. N. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Springer Science + Business Media, 1933.

[7] A. N. Kolmogorov. *Foundations of the theory of probability.* Chelsea Publishing Company, 1956.

[8] Y. Y. Melamed and M. Lin. Principle of Sufficient Reason. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, spring 2020 edition, 2020.

[9] I. C. Percival. *Quantum State Diffusion.* Cambridge University Press, 2003.

[10] A. Peres. *Quantum theory: concepts and methods*, volume 57 of *Fundamental Theories of Physics.* Springer, 1995.

[11] J. von Neumann. *Mathematische Grundlagen der Quantenmechanik. (English translation "The Mathematical Foundations of Quantum Mechanics" Princeton University Press 1955).* Grundlehren der Mathematischen Wissenschaften. Springer Science + Business Media, 2nd edition, 1981.