QUANTUM INEQUALITIES AND SEQUENTIAL MEASUREMENTS

Bernard Candelpergher[†]

Université de Nice — Sophia-Antipolis (UFR Sciences)
Laboratoire Jean Alexandre Dieudonné — CNRS (UMR6621)
Parc Valrose, 28 avenue Valrose, 06108 Nice Cedex 2, France

Thierry Grandou[‡], Jacques L. Rubin[§]

Université de Nice — Sophia-Antipolis (UFR Sciences)
Institut du Non-Linéaire de Nice — CNRS (UMR6618)
1361 route des Lucioles, 06560 Valbonne, France

(Received August 31, 2010; revised version received December 1, 2010; final version received January 10, 2011)

In this article, the peculiar context of sequential measurements is chosen in order to analyze the quantum specificity in the two most famous examples of Heisenberg and Bell inequalities: Results are found at some interesting variance with customary textbook materials, where the context of initial state re-initialization is described. A key-point of the analysis is the possibility of defining Joint Probability Distributions for sequential random variables associated to quantum operators. Within the sequential context, it is shown that Joint Probability Distributions can be defined in situations where not all of the quantum operators (corresponding to random variables) do commute two by two.

DOI:10.5506/APhysPolB.42.209
PACS numbers: 03.65.Ud, 03.65.Ta

1. Introduction

In line with the basics of Copenhagen interpretation, it is known that Quantum Mechanics retain a dependence on experimental devices, protocols and environment and is henceforth, somewhat *contextual.* In this article the

[†] bernard.candelpergher@unice.fr
[‡] thierry.grandou@inln.cnrs.fr
[§] jacques.rubin@inln.cnrs.fr

quantum specificity is examined within the context of sequential measurement processes, and two among the most famous inequalities of Quantum Mechanics are analyzed, Bell's and Heisenberg's inequalities.

The article is organized as follows. After a brief reminder of probability theory in Section 2, the quantum formalism is introduced in Section 3, and within it, the central definition of (quantum) random variables (rvs, for short) associated to sequential measurements of observables, themselves associated to quantum operators that do not all necessarily commute two by two, [1].

On states and/or operators, sequences are generated by the action of an *intertwinning operator* whose properties are crucial to our concern and are dealt with in Section 4.

In Section 5, the intertwinning operator allows us to define the (quantum) rvs associated to sequential measurements, and analyze their first properties, with much attention paid to the possibility of defining Joint Probability Distributions for rvs corresponding to sequential measurements.

A notion of *statistical insensivity* can be introduced in relation to the fact that, if two observables are measured sequentially, then their probability distributions may come out unaffected by the process. Likewise, rvs may or may not be *statistically independent*. The reciprocal relations of these properties, envisaged in connection to the commutativity or non-commutativity of the corresponding quantum operators, are analyzed in Sections 6 and are summarized in a diagram displaying the implications and non-implications between them.

The elements gathered in these previous sections allow us to study the cases of Heisenberg and Bell's inequalities in a context of sequential measurements. This is done in Section 7 for the former and in Section 8 for the latter, whereas Section 9 presents our conclusions.

Throughout the article, several simple examples are proposed to illustrate the content of notions that could otherwise appear a bit formal, and it is hoped that this will provide the text with some pedagogical value either.

2. Statistical formalism of probability theory

All the random variables that we consider are supposed to take only a finite number of values. The probability distribution of a random variable (rv) X is defined by:

— its possible values: $x_1, \ldots, x_n$,

— and associated probabilities: $p(X = x_1), \ldots, p(X = x_n)$.

These probabilities are positive numbers such that $\sum_i p(X = x_i) = 1$ and are interpreted as the frequencies of instance of the respective values over a large number of experimental realizations.

The mean value of the rv X is, by definition, the number

$$E(X) = \sum_i x_i \, p(X = x_i) \, .$$

Definition 1 *Let X,Y be two random variables with values $x_1, \ldots, x_n$, and $y_1, \ldots, y_m$. We say that (X, Y) satisfies the criterium of* **sequential measurements** *if, for each instance of the experiment, a measurement of X followed by a measurement of Y can be performed.*

This allows us to define the number $N(x_i, y_j, n)$ of instances of $X = x_i$ and $Y = y_j$ as the experiment is repeated n times.

Definition 2 *Let X,Y be two random variables with values $x_1, \ldots, x_n$, and $y_1, \ldots, y_m$. A* **compatible probability distribution** *is a two dimensional probability distribution*

$(p_{i,j}), i = 1, \ldots, n, j = 1, \ldots, m$ that satisfies the "marginal" laws:

$$\sum_i p_{i,j} = p(Y = y_j) \, , \qquad \sum_j p_{i,j} = p(X = x_i) \, .$$

Remark Note that given the probability distributions of X and Y there is always an infinite number of compatible probability distribution. For example, if

$$\begin{aligned} p(X &= 1) = p \, , \qquad p(X = 0) = 1 - p \, , \\ p(Y &= 1) = q \, , \qquad p(Y = 0) = 1 - q \, . \end{aligned}$$

The compatible probability distributions $(p_{i,j})$ are given by arrays of numbers $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$, such that

$$\begin{aligned} a &= \alpha \, , \\ b &= p - \alpha \, , \\ c &= q - \alpha \, , \\ d &= 1 + \alpha - p - q \, , \end{aligned}$$

with $\max(0, p + q - 1) \leq \alpha \leq \min(p, q)$. In particular for $\alpha = pq$, we get the probability distribution

$$p_{i,j} = p(X = x_i) p(Y = y_j) \, ,$$

which always defines a (trivial) compatible probability distribution.

Remark In the Kolmogorov formalism of Probability Theory, the random variables X, Y are defined as functions on a set Ω endowed with a probability measure p. The probability distributions of X and Y are defined by

$$\begin{aligned} p_X(x_i) &= p\left(\{\omega \in \Omega \text{ such that } X(\omega) = x_i\}\right) , \\ p_Y(y_j) &= p\left(\{\omega \in \Omega \text{ such that } Y(\omega) = y_j\}\right) . \end{aligned}$$

The joint probability distributions of X and Y is defined by

$$p_{X,Y}(x_i, y_j) = p\left(\{\omega \in \Omega \text{ such that } X(\omega) = x_i \text{and} Y(\omega) = y_j\}\right) .$$

Within these definitions it is straightforward to verify that this joint distribution is clearly compatible with the marginals.

Definition 3 *Joint probability distribution. Let X, Y be two random variables with values $x_1, \dots, x_n$, and $y_1, \dots, y_m$. We say that (X, Y) admits the* **joint probability distribution (JPD)**$(p_{i,j}), i = 1, \dots, n, j = 1, \dots, m$, *if:*

— (X, Y) *satisfies the sequential measurement criterium,*

— $(p_{i,j})$ *is a two dimensional compatible probability distribution, such that*

$$p_{i,j} = \lim_{n\to\infty} \frac{N(x_i, y_j, n)}{n} .$$

We note $p_{X,Y}$ this JPD:

$$p_{X,Y}(x_i, y_j) = p(X = x_i,\, Y = y_j) = p_{i,j} .$$

These definitions allow one to test in a practical (statistical) point of view, if a given two dimensional compatible probability distribution $(p_{i,j})$ is a statistically relevant JPD for (X, Y).

Remark Sum and product of two rvs. If X and Y are two random variables then the sum $X + Y$ or product XY or any function of them, $f(X, Y)$, is not necessarily defined as a random variable. For example, if the values of the variable $X + Y$ are the values of $z \in \{x_i + y_j$, such that, $i = 1, 2, \dots, n,\ j = 1, 2, \dots, m\}$, still, the probability of these values must be defined. If there exists a JPD of (X, Y), (p_{ij}), then, these probabilities are given by

$$p(X + Y = z) = \sum_{\substack{x_i + y_j = z \\ i, j}} p_{(X,Y)}(x_i, y_j) . \tag{1}$$

In this situation, the relation $E(X + Y) = E(X) + E(Y)$ is easily proven to hold. In the same way, the random variable XY can be defined as the rv with values the $z \in \{x_i\, y_j$ such that $i = 1, \ldots, n$ and $j = 1, \ldots, m\}$, and associated probabilities, the numbers

$$p(X\,Y = z) = \sum_{\substack{x_i\, y_j = z \\ i,\, j}} p(X = x_i,\, Y = y_j)\,. \tag{2}$$

The relation $E(X\,Y) = E(X)\,E(Y)$ does not hold necessarily.

Definition 4 *Statistical independence. If X and Y are two random variables, we say that X and Y are* **independent** *if, for all i and all j, their JPD satisfies*

$$p(X = x_i,\, Y = y_j) = p(X = x_i)\, p(Y = y_j)\,. \tag{3}$$

If X and Y are two independent random variables, the relation $E(X\,Y) = E(X)\,E(Y)$ holds trivially.

3. Quantum formalism

The formalism of Quantum Mechanics can be developed in terms of the following basic objects and postulates (evolution, which would complete the usual quantum mechanical description, will not be dealt with in the sequel).

Statistical Operator. Over $\mathbb{E}$, a $\mathbb{C}$–Hilbert space endowed with the inner product $\langle\;|\;\rangle$, and $\mathcal{H}(\mathbb{E})$, the set of Hermitian operators on $\mathbb{E}$, the system is described by a statistical operator (or density matrix), $\rho \in \mathcal{H}(\mathbb{E})$, satisfying:

— $\mathrm{Tr}(\rho) = 1$,

— ρ positive, *i.e.*, $\forall \psi \in \mathbb{E}$, $\langle \psi | \rho\, \psi \rangle \geq 0$.

Postulate 1 For any observable $A \in \mathcal{H}(\mathbb{E})$, one has $A = \sum_i a_i\, P_{A_i}$, where $a_i \in \mathbb{R}$, and where the P_{A_i} are projectors on the proper subspaces associated to the $a_i s$, including the possible vanishing eigenvalue (*i.e.*, the relation $\sum_i P_{A_i} = \mathbb{1}$ is satisfied in the definition above for A). Then, to any state ρ and any observable A one can associate the rv $X_A(\rho)$ representing the result of the measure of A in state ρ.

The random variable $X_A(\rho)$ is therefore defined by:

— the a_i, as its possible values, with

— associated probabilities, [2], the $p(X_A(\rho) = a_i) = \mathrm{Tr}(\rho\, P_{A_i})$.

Postulate 2 Wave Packet Reduction postulate. Though this postulate is no longer considered as such nowadays, and seems to be reducible to a more intuitive statement [3], it can still be used as a convenient way to put things [4].

A quantum system is described by the statistical state ρ. An observable corresponding to the operator A, is measured on it. Then the system's statistical state is $\mathit{\Gamma}_A(\rho)$, where $\mathit{\Gamma}_A$ is the "intertwining operator" defined on $\mathcal{H}(\mathbb{E})$ by

$$B \longmapsto \mathit{\Gamma}_A(B) = \sum_i P_{A_i}\, B\, P_{A_i}\,, \tag{4}$$

where $\sum_i P_{A_i} = \mathbb{1}$.

The intertwinning operator. In this section some properties of the intertwinning operator just defined, (4), are given, that will be helpful in the sequel.

On $\mathcal{H}(\mathbb{E})$ we define the scalar product of A and B, by $(A,B) = \mathrm{Tr}(AB)$. The operator $\mathit{\Gamma}_A$ from $\mathcal{H}(\mathbb{E})$ into $\mathcal{H}(\mathbb{E})$ is self adjoint,

$$(\mathit{\Gamma}_A B, C) = (B, \mathit{\Gamma}_A C)\,,$$

and the following properties can be verified easily:

Properties

(1) If $[A,B] = 0$, then $\mathit{\Gamma}_A B = B$.

(2) For all C, one has
$$\mathit{\Gamma}_A(BC) = B(\mathit{\Gamma}_A C)\,.$$

(3) $\mathit{\Gamma}_A B$ commutes with A.

(4) $\mathit{\Gamma}_A(BA) = A\,\mathit{\Gamma}_A B = \mathit{\Gamma}_A(AB)$.

(5) Eventually, one can quote the property that if $B = \rho$ is a density matrix, then, so is also $\mathit{\Gamma}_A(\rho)$.

(6) If the eigenspaces A_i have dimension 1, then
$$\mathit{\Gamma}_A B = \sum_i \mathrm{Tr}(P_{A_i} B)\, P_{A_i}\,,$$
in view of the fact that for all operator B and all subspace $F \in \mathbb{E}$, if $\dim(F) = 1$, then $P_F B P_F = \mathrm{Tr}(B P_F)\, P_F$.

4. Succession of measurements

Let us define $X_{B\leftarrow A}(\rho)$, a new random variable, associated to a sequence where a measure of an observable B, in a given quantum system, is performed after another observable A has been measured on the same system (here, "after" is to be understood in a non-relativistic acceptation).

Definition 5 *Let A, $B \in \mathcal{H}(\mathbb{E})$. One defines the rv $X_{B\leftarrow A}(\rho)$ by*

$$X_{B\leftarrow A}(\rho) = X_B(\Gamma_A(\rho))\,. \tag{5}$$

Thus the random variable $X_{B\leftarrow A}(\rho)$, has the probability distribution:

— *b_1, …, b_n, the eigenvalues of B,*

— *with associated probabilities $p(X_{B\leftarrow A}(\rho) = b_j) = \mathrm{Tr}\left(\Gamma_A(\rho)\, P_{B_j}\right)$.*

Remark One has

$$E(X_B(\Gamma_A(\rho)) = (\Gamma_A(\rho)|B) = (\rho|\Gamma_A B) = E(X_{\Gamma_A B}(\rho))\,.$$

However, $\Gamma_A B$ does not necessarily possess the b_j as its eigenvalues, so that in the general case it matters to realize that

$$X_B(\Gamma_A(\rho)) \neq X_{\Gamma_A B}(\rho)\,.$$

Properties

(1) Note that if $B = A$, then

$$p\left(X_{A\leftarrow A}(\rho) = a_j\right) = \mathrm{Tr}\left(\Gamma_A(\rho)\, P_{A_j}\right) = \mathrm{Tr}\left(\rho\, P_{A_j}\right) = p\left(X_A(\rho) = a_j\right)\,,$$

thus we have

$$X_{A\leftarrow A}(\rho) = X_A(\rho)\,. \tag{6}$$

(2) Also,

$$E(X_{B\leftarrow A}(\rho)) = E(X_B(\Gamma_A(\rho))) = (\Gamma_A(\rho), B) = (\rho, \Gamma_A B)\,. \tag{7}$$

Theorem 1 *The family of numbers*

$$p_{B\leftarrow A,\rho}(a_i, b_k) = \mathrm{Tr}\left(P_{A_i}\,\rho\, P_{A_i} P_{B_k}\right)$$

satisfies the conditions

— $\sum\limits_k p_{B\leftarrow A,\rho}(a_i, b_k) = p(X_A(\rho) = a_i)$,

— $\sum\limits_i p_{B\leftarrow A,\rho}(a_i, b_k) = p(X_{B\leftarrow A}(\rho) = b_k)$.

Proof. Obvious, by partial summations.

Postulate 3 (Wigner formula, [5]); The couple of random variables $(X_A(\rho), X_{B\leftarrow A}(\rho))$, admits the JPD whose array of numbers is given by the Wigner formula [5]

$$p(X_A(\rho) = a_i, X_{B\leftarrow A}(\rho) = b_k) = \mathrm{Tr}\left(P_{A_i}\,\rho\, P_{A_i} P_{B_k}\right)\,. \tag{8}$$

Theorem 2 *The JPD of* $(X_A(\rho), X_{B\leftarrow A}(\rho))$*, allows us to define the random variables* $X_A(\rho) + X_{B\leftarrow A}(\rho)$ *and* $X_A(\rho)\, X_{B\leftarrow A}(\rho)$*, and one has*

$$\begin{aligned} E(X_A(\rho) + X_{B\leftarrow A}(\rho)) &= (\rho, A + \Gamma_A(B))\,, \\ E(X_A(\rho)\, X_{B\leftarrow A}(\rho)) &= (\Gamma_A(\rho), AB)\,. \end{aligned}$$

The proof is obvious.

Remark If, according to what precedes, the rvs $X_A(\rho) + X_{B\leftarrow A}(\rho)$ and $X_A(\rho)\, X_{B\leftarrow A}(\rho)$ are defined, such is not necessarily the case of the variable $X_A(\rho) + X_B(\rho)$ because nothing guarantees that the couple $(X_A(\rho),\, X_B(\rho))$ admits a JPD. As a matter of fact, $E(X_A(\rho)) + E(X_B(\rho))$ is defined, whereas this may be not so for $E(X_A(\rho) + X_B(\rho))$. Generalized to a Bell's inequality, a similar remark will be made in Subsection 8.2.

5. Statistical insensitivity

Definition 6 *In a state* ρ*,* B *will be said to be "statistically insensitive to"* A*, if the rvs* $X_{B\leftarrow A}(\rho)$ *and* $X_B(\rho)$ *have the same probability distribution. That is, in a given state* ρ*,* B *is statistically insensitive to* A*, if*

$$\forall\, j,\ \mathrm{Tr}\left(\Gamma_A(\rho) P_{B_j}\right) = \mathrm{Tr}\left(\rho\, P_{B_j}\right)\,. \tag{9}$$

We note this property, $X_{B\leftarrow A}(\rho) \sim X_B(\rho)$*.*

In other words, a measure of A in state ρ does not affect the statistical behavior of X_B as B is measured after A.

Taking A to be a projector, then

$$\mathrm{Tr}\left(\Gamma_A(\rho) P_{B_j}\right) = \mathrm{Tr}\left(P_{A_0} \rho P_{A_0} P_{B_j}\right) + \mathrm{Tr}\left(P_{A_1} \rho P_{A_1} P_{B_j}\right)\,,$$

and

$$\mathrm{Tr}\left(\rho P_{B_j}\right) = \mathrm{Tr}\left((P_{A_0} + P_{A_1})\rho(P_{A_0} + P_{A_1}) P_{B_j}\right)\,.$$

Thus the condition that B is statistically insensitive to A gives

$$\mathrm{Re}\left(\mathrm{Tr}\left(P_{A_0} \rho P_{A_1} P_{B_j}\right)\right) = 0\,,$$

which is nothing else than Griffiths histories consistency relations [6].

Definition 7 *In a state* ρ*,* A *and* B *are "mutually insensitive" if* $X_{B\leftarrow A}(\rho) \sim X_B(\rho)$ *and* $X_{A\leftarrow B}(\rho) \sim X_A(\rho)$*. That is, if the following conditions are met:*

— $\forall\, k,\ \mathrm{Tr}(\Gamma_A(\rho)\, P_{B_k}) = \mathrm{Tr}\,(\rho\, P_{B_k})$,

— $\forall\, j,\ \mathrm{Tr}\big(\Gamma_B(\rho)\, P_{A_j}\big) = \mathrm{Tr}\left(\rho\, P_{A_j}\right)$.

At this level, one may note that the more intuitive basis from which Postulate 2 can be deduced can be stated as the condition $X_{A\leftarrow A}(\rho) = X_A(\rho)$ (6), when two measurements of A are performed at very close instants in time [3].

As a first result, it is elementary to check that the following result holds.

Theorem 3 *If* $[A,B]=0$, *then* $\forall\,\rho$

$$\begin{aligned} X_{B\leftarrow A}(\rho) &\sim X_B(\rho)\,, \\ X_{A\leftarrow B}(\rho) &\sim X_A(\rho) \end{aligned}$$

that is, A and B are mutually insensitive in all states ρ. The JPDs of the two couples of rvs $(X_A(\rho), X_{B\leftarrow A}(\rho))$ and $(X_{A\leftarrow B}(\rho), X_B(\rho))$ are the same, and we have

$$\begin{aligned} p(X_A(\rho) &= a_i, X_{B\leftarrow A}(\rho) = b_k) = p(X_{A\leftarrow B}(\rho) = a_i, X_B(\rho) = b_k) \\ &= \mathrm{Tr}(P_{A_i}\,\rho\,P_{B_k})\,. \end{aligned}$$

Proof: If $[A,B]=0$, then $\forall\,j$, $[A, P_{B_j}] = 0$ and therefore, $\mathit{\Gamma}_A(P_{B_j}) = P_{B_j}$. So,

$$\begin{aligned} p(X_{B\leftarrow A}(\rho) &= b_j) = (\mathit{\Gamma}_A(\rho),\, P_{B_j}) = (\rho,\, \mathit{\Gamma}_A(P_{B_j})) \\ &= (\rho,\, P_{B_j}) = p(X_B(\rho) = b_j)\,, \end{aligned}$$

and likewise for $X_{A\leftarrow B}(\rho) = X_A(\rho)$. Moreover, we have

$$\begin{aligned} &p(X_A(\rho) = a_i, X_{B\leftarrow A}(\rho) = b_k) \\ &= \mathrm{Tr}(P_{A_i}\,\rho\,P_{A_i}P_{B_k}) = \mathrm{Tr}(P_{A_i}\,\rho\,P_{B_k}P_{A_i}) = \mathrm{Tr}(P_{A_i}\,\rho\,P_{B_k})\,, \\ &p(X_{A\leftarrow B}(\rho) = a_i, X_B(\rho) = b_k) \\ &= \mathrm{Tr}(\,\rho\,P_{B_k}P_{A_i}P_{B_k}) = \mathrm{Tr}(\,\rho\,P_{B_k}P_{A_i}) = \mathrm{Tr}(P_{A_i}\,\rho\,P_{B_k})\,. \end{aligned}$$

End of proof.

Remark If $[A,B]\neq 0$, the two couples of rvs $(X_A(\rho), X_{B\leftarrow A}(\rho)\,)$ and $(X_{A\leftarrow B}(\rho), X_B(\rho)\,)$ have the JPDs that are not necessarily the same. This does not imply either, that the couple $(X_A(\rho), X_B(\rho))$ possesses a JPD.

When $[A,B]=0$, the JPDs (Theorem 3) are not affected by the order according to which measurements are performed, and one may therefore decide that the couple of rvs $(X_A(\rho), X_B(\rho))$ admits the JPD

$$p(X_A(\rho) = a_i, X_B(\rho) = b_k) = \mathrm{Tr}(P_{A_i}\,\rho\,P_{B_k})\,.$$

For this JPD, one gets

$$X_A(\rho) + X_B(\rho) = X_{A+B}(\rho)\,, \qquad X_A(\rho)\,X_B(\rho) = X_{AB}(\rho)\,.$$

In this situation, the remarkable fact is that a simple mapping of the algebra generated by a complete set of commuting observables (CSCO), into the corresponding rvs set, is realized.

6. Commutativity, insensitivity and statistical independence

It is interesting to compare the properties of commutativity, statistical independence and insensitivity.

6.1. Commutativity vs. insensitivity

1. Theorem 3 shows that commutativity implies insensitivity.
2. Insensitivity does not imply commutativity: in a state like $\rho = \frac{1}{n}\mathbb{1}$, for example, all observables are mutually insensitive.

6.2. Insensivity vs. independence

Having a JPD for the couple of random variables $(X_A(\rho), X_{B\leftarrow A}(\rho))$ in view of Postulate 3, one can analyze their statistical independence.

1. Indeed, one can observe that random variables can be statistically independent in a context where statistical insensitivity does not hold (and thus commutativity does not hold). In other words, a measure of A modifies the probability distribution of X_B, while $X_A(\rho)$ and $X_{B\leftarrow A}(\rho)$ display no correlation.

   Example: Consider $E = \mathbb{C}^2$ and ρ, the state

   $$\rho = \begin{pmatrix} \alpha & \beta \\ \beta & 1-\alpha \end{pmatrix}$$

   at $0 < \alpha^2 + \beta^2 \leq \alpha < 1$, and let A be the observable

   $$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

   with eigenvalues $a_1 = 1$ and $a_{-1} = -1$. The probability distribution of $X_A(\rho)$ reads as

   $$p(X_A(\rho) = 1) = \beta + \tfrac{1}{2}, \qquad p(X_A(\rho) = -1) = -\beta + \tfrac{1}{2}.$$

   Consider also

   $$B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

   whose eigenvalues are $b_1 = 1$ and $b_{-1} = -1$. The probability distribution of $X_B(\rho)$ reads as

   $$p(X_B(\rho) = 1) = \alpha, \qquad p(X_B(\rho) = -1) = 1 - \alpha.$$

One has

$$\Gamma_A(\rho) = \begin{pmatrix} \frac{1}{2} & \beta \\ \beta & \frac{1}{2} \end{pmatrix},$$

so that the probability distribution of $X_B(\Gamma_A(\rho))$ is given by

$$p(X_B(\Gamma_A(\rho)) = 1) = p(X_B(\Gamma_A(\rho)) = -1) = \tfrac{1}{2}.$$

The natural JPD for $X_A(\rho)$ and $X_B(\Gamma_A(\rho))$ reads therefore as

$$\begin{aligned} &p((X_A(\rho), X_{B\leftarrow A}(\rho)) = (1,1)) \\ &= p((X_A(\rho), X_{B\leftarrow A}(\rho)) = (1,-1)) = \tfrac{1}{4} + \tfrac{1}{2}\beta, \\ &p((X_A(\rho), X_{B\leftarrow A}(\rho)) = (-1,-1)) \\ &= p((X_A(\rho), X_{B\leftarrow A}(\rho)) = (-1,1)) = \tfrac{1}{4} - \tfrac{1}{2}\beta. \end{aligned}$$

One then observes that $X_A(\rho)$ and $X_B(\Gamma_A(\rho))$ are statistically independent. Still, if $\alpha < 1/2$ the rvs $X_B(\Gamma_A(\rho))$ and $X_B(\rho)$ *have not* the same distribution.

2. Likewise, insensitivity does not imply independence either. This can be seen on the following example.

   Example: $E = \mathbb{C}^2$ and

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad A = \begin{pmatrix} a_1 & 0 \\ 0 & a_2 \end{pmatrix}, \qquad B = \begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix}.$$

   One has $AB = BA$ and

$$p(X_A(\rho) = a_1) = \mathrm{Tr}(\rho\, P_{A_1}) = \tfrac{1}{2}, \quad p(X_B(\rho) = b_1) = \mathrm{Tr}(\rho\, P_{B_1}) = \tfrac{1}{2}.$$

   Now, it can be checked that

$$\begin{aligned} &p(X_A(\rho) = a_1,\, X_B(\rho) = b_1) = \mathrm{Tr}(\rho\, P_{A_1} P_{B_1}) \\ &= \tfrac{1}{2} \neq \mathrm{Tr}(\rho\, P_{A_1})\mathrm{Tr}(\rho\, P_{B_1}). \end{aligned}$$

6.3. Commutativity vs. independence

1. On the other hand, if one takes $B = A$, then $[A, B] = 0$, but $X_A(\rho)$ and $X_A(\Gamma_A(\rho)) = X_A(\rho)$ are not statistically independent if $X_A(\rho)$ is not a constant rv, and this illustrates the fact that commutativity does not imply independence.

2. We have just seen in Subsec. 6.2 that one can have independence without insensitivity, thus independence does not imply commutativity.

The following diagram summarizes some of the implications we have just dealt with.

$$\begin{array}{ccc}
[A,B]=0 & & \\
 & \not\Longleftrightarrow & \\
\Big\Downarrow & & \begin{array}{c} X_A(\rho) \text{ and } X_{B\leftarrow A}(\rho) \\ \text{are independent} \end{array} \\
 & \not\Longleftrightarrow & \\
\begin{array}{c} A \text{ and } B \text{ mutually} \\ \text{insensitive in state } \rho \end{array} & &
\end{array}$$

Remark There is an important case where commutativity implies statistical independence. Consider a system whose space of states is a tensorial product $\mathbb{E} \otimes \mathbb{F}$. To A, an observable on $\mathbb{E}$, is associated $A \otimes I$ on $\mathbb{E} \otimes \mathbb{F}$, and likewise, to B an observable on $\mathbb{F}$ is associated an observable $I \otimes B$ on $\mathbb{E} \otimes \mathbb{F}$. The operators $A \otimes I$ and $I \otimes B$ commute, and indeed

$$X_{A\otimes B} = X_{(A\otimes I)(I\otimes B)} = X_{(A\otimes I)} X_{(I\otimes B)} \,. \tag{10}$$

Since the two above operators commute, $A \otimes I$ and $I \otimes B$ are mutually insensitive, and the rvs $X_{A\otimes I}$ and $X_{I\otimes B}$ admit as a natural JPD, the probability distribution

$$\begin{aligned} p(X_{A\otimes I}(\rho) = a_i \text{ and } X_{I\otimes B}(\rho) = b_j) \;&=\; \mathrm{Tr}\left(\rho(p_{A_i} \otimes I)(I \otimes p_{B_j})\right) \\ &=\; \mathrm{Tr}\left(\rho(p_{A_i} \otimes p_{B_j})\right) \,. \end{aligned} \tag{11}$$

If the system is in a tensorial product state such as $\rho = \alpha \otimes \beta$, then the rvs $X_{A\otimes I}(\alpha \otimes \beta)$ and $X_{I\otimes B}(\alpha \otimes \beta)$ have the same probability distributions as $X_A(\alpha)$ and $X_B(\beta)$, and the following relation holds

$$\begin{aligned} & p(X_{A\otimes I}(\alpha \otimes \beta) = a_i \text{ and } X_{I\otimes B}\,(\alpha \otimes \beta) {=} b_j) \\ & {=} \mathrm{Tr}\left(\alpha \otimes \beta(p_{A_i} \otimes p_{B_j})\right) {=} \mathrm{Tr}(\alpha(p_{A_i}))\mathrm{Tr}\left(\beta(p_{B_j})\right) \,, \end{aligned} \tag{12}$$

thus, the rvs $X_{A\otimes I}(\alpha \otimes \beta)$ and $X_{I\otimes B}(\alpha \otimes \beta)$ are statistically independent and therefore,

$$E(X_{A\otimes B}(\alpha \otimes \beta)) = E(X_A(\alpha))E(X_B(\beta)) \,. \tag{13}$$

7. Heisenberg inequalities

The notions and properties gathered so far, allow one to analyze some famous quantum inequalities in the peculiar context of sequential measurements. This will first be done in the case of the Heisenberg inequality.

Let X and Y be two rvs endowed with a JPD, and let us define $\widetilde{X} = X - E(X)$ and $\widetilde{Y} = Y - E(Y)$. The following *Cauchy–Schwarz inequality* (in the space $L^2(\mathbb{R}^2, d\mu)$, where μ is the discrete measure associated to the JPD of X and Y)

$$\left(E\left(\widetilde{X}\widetilde{Y}\right)\right)^2 \leq E\left(\left(\widetilde{X}\right)^2\right) E\left(\left(\widetilde{Y}\right)^2\right), \tag{14}$$

holds since the dispersions of X and Y are defined by the relations

$$\sigma(X) = \sqrt{E((X - E(X))^2)}\,, \qquad \sigma(Y) = \sqrt{E((Y - E(Y))^2)}\,. \tag{15}$$

A *classical dispersion's inequality* can therefore be proposed under the form

$$\sigma(X)\sigma(Y) \geq |E(XY) - E(X)E(Y)|\,. \tag{16}$$

Note that in the quantum case, one cannot apply this inequality to the rvs $X_A(\rho)$ and $X_B(\rho)$ since, in general, one does not know if there is a JPD for these rvs. It is well known, though, that this inequality has a famous analogue in quantum mechanics.

7.1. Quantum dispersion inequality

For any statistical operator ρ, and any A and B in $\mathcal{L}(\mathbb{E})$, let us define the product

$$(A|B)_\rho = \mathrm{Tr}(A^\star \rho B)\,, \tag{17}$$

where $A^\star$ is the standard hermitian conjugate of A.

Lemma 1 *The above product defines a hermitian scalar product on* $\mathcal{L}(\mathbb{E})$.

Proof: First we can prove that for any $C \in \mathcal{L}(\mathbb{E})$, if we have $(C|C)_\rho = 0$, then, $C = 0$. Now, the operator $CC^\star$ is hermitian and positive and thus the rv $X_{CC^\star}(\rho)$ has positive values. Therefore,

$$E(X_{CC^\star}(\rho)) = \mathrm{Tr}(\rho\, CC^\star) \geq 0\,,$$

and if $\mathrm{Tr}(\rho\, CC^\star) = 0$, then $X_{CC^\star}(\rho) = 0$, and thus one has $CC^\star = 0$, that is $C = 0$. Furthermore, the property $(B|A)_\rho = \overline{(A|B)_\rho}$ is a mere consequence of both the cyclicity of the Trace and of the hermiticity of ρ. End of proof.

For A an observable, the dispersion of the corresponding rv $X_A(\rho)$ is defined to be

$$\sigma(X_A(\rho)) = \sqrt{\mathrm{Var}(X_A(\rho))} = \sqrt{E((X_A(\rho))^2) - (E(X_A(\rho)))^2}\,. \tag{18}$$

Setting, $\widetilde{A_\rho} = A - E((X_A(\rho)) = A - \langle A\rangle_\rho$, one gets

$$\sigma(X_A(\rho)) = \sqrt{\mathrm{Tr}(\rho(\widetilde{A_\rho})^2)} = ||\widetilde{A_\rho}||_\rho\,. \tag{19}$$

Then, in $\mathcal{H}(\mathbb{E})$, the Cauchy–Schwarz inequality

$$\|\widetilde{A_\rho}\|_\rho\,\|\widetilde{B_\rho}\|_\rho \geq \left|(\widetilde{A_\rho}|\widetilde{B_\rho})_\rho\right|, \tag{20}$$

leads to the *quantum dispersion inequality*

$$\sigma(X_A(\rho))\,\sigma(X_B(\rho)) \geq \left|\mathrm{Tr}(\rho\,A\,B) - \langle A\rangle_\rho \langle B\rangle_\rho\right|. \tag{21}$$

Out of the above inequality, two *symmetric* and *antisymmetric* inequalities follow:

— A symmetric inequality

$$\|\widetilde{A}_\rho\|_\rho\,\|\widetilde{B}_\rho\|_\rho \geq \left|Re(\widetilde{A}_\rho\middle|\widetilde{B}_\rho)_\rho\right| = \left|\tfrac{1}{2}\,\mathrm{Tr}(\rho\,(AB+BA)) - \langle A\rangle_\rho \langle B\rangle_\rho\right|, \tag{22}$$

or,

$$\sigma(X_A(\rho))\,\sigma(X_B(\rho)) \geq \left|\tfrac{1}{2}\,\mathrm{Tr}(\rho\,(AB+BA)) - \langle A\rangle_\rho \langle B\rangle_\rho\right|, \tag{23}$$

— An antisymmetric inequality

$$\|\widetilde{A}_\rho\|_\rho\,\|\widetilde{B}_\rho\|_\rho \geq \left|Im(\widetilde{A}_\rho|\widetilde{B}_\rho)_\rho\right| = \tfrac{1}{2}\left|\mathrm{Tr}(\rho\,[A,B])\right| \tag{24}$$

or, the generalized Heisenberg inequality

$$\sigma(X_A(\rho))\,\sigma(X_B(\rho)) \geq \tfrac{1}{2}\left|\mathrm{Tr}(\rho\,[A,B])\right|, \tag{25}$$

whereof the standard Heisenberg inequality

$$\sigma(X_A(\rho))\,\sigma(X_B(\rho)) \geq \tfrac{1}{2}\,|\langle\psi|[A,B]\psi\rangle| \tag{26}$$

results in the case of a pure state $\rho = |\psi\rangle\langle\psi|$.

7.2. Application to sequential measurements

Ordinarily, the Heisenberg inequality refers to the dispersions of the rvs $X_A(\rho)$ and $X_B(\rho)$, that is, implicitly, to an experimental protocol in which the system is, at each instance, re-prepared in its initial state ρ, [7,8]. And so a natural issue is the one related to the dispersions of rvs $X_A(\rho)$ and $X_{B\leftarrow A}(\rho)$, corresponding to another experimental protocol where such a re-initialization of the system is precluded. In the case of sequential measurements, the rvs $X_A(\rho)$ and $X_{B\leftarrow A}(\rho)$ have a JPD, in view of Postulate 3, so that, in this situation, the classical dispersion inequality (16) also applies to the dispersions of the *quantum* rvs

$$\sigma(X_A(\rho))\,\sigma(X_{B\leftarrow A}(\rho)) \geq |E(X_A(\rho)\,X_{B\leftarrow A}(\rho)) - E(X_A(\rho))\,E(X_{B\leftarrow A}(\rho))|\,,$$

which gives

$$\sigma(X_A(\rho))\,\sigma(X_{B\leftarrow A}(\rho)) \geq |\mathrm{Tr}(\mathit{\Gamma}_A(\rho)\,A\,B) - \langle A\rangle_\rho\,\langle B\rangle_{\mathit{\Gamma}_A(\rho)}|\,,$$

that is, remarkably enough, nothing but the quantum dispersion inequality itself, (21), because of (6). In a state that would simultaneously be an eigenvector of both A and $\mathit{\Gamma}_A(B)$, we note that the above right hand side is zero [9]. Writing

$$A\,B = \tfrac{1}{2}\,(A\,B - B\,A) + \tfrac{1}{2}\,(A\,B + B\,A)\,,$$

we can decompose the preceding inequality into an analogue of symmetric and antisymmetric parts, the antisymmetric part reading,

$$\sigma(X_A(\rho))\,\sigma(X_{B\leftarrow A}(\rho)) \geq \tfrac{1}{2}\left|\mathrm{Tr}(\mathit{\Gamma}_A(\rho)(AB - BA))\right|.$$

This inequality is nothing else than the Heisenberg inequality (25) applied to the state $\mathit{\Gamma}_A(\rho)$. However, this antisymmetric part vanishes because one has

$$\begin{aligned}\tfrac{1}{2}\mathrm{Tr}(\mathit{\Gamma}_A(\rho)(AB-BA)) &= \tfrac{1}{2}\sum_i \mathrm{Tr}(P_{A_i}\rho P_{A_i}A\,B) - \tfrac{1}{2}\sum_i \mathrm{Tr}(AP_{A_i}\rho P_{A_i}B)\\ &= \tfrac{1}{2}\sum_i a_i\mathrm{Tr}(P_{A_i}\rho P_{A_i}B) - \tfrac{1}{2}\sum_i a_i\mathrm{Tr}(P_{A_i}\rho P_{A_i}B)\\ &= 0\,,\end{aligned} \tag{27}$$

so that only the trivial inequality results in $\sigma(X_A(\rho))\,\sigma(X_{B\leftarrow A}(\rho)) \geq 0$. As it stands, the above inequality does not teach much indeed. However, it may suggest that sequential measurements are able to somewhat reduce the rvs dispersions. The following example provides an illustration of a fact that could be of some interest in view of the results recently put forth by Katz and *et al.* [10].

Example: We consider the Hilbert space $\mathbb{C}^2 \otimes \mathbb{C}^2$, spanned by the basis vectors $(0 \otimes 0)$, $(0 \otimes 1)$, $(1 \otimes 0)$, $(1 \otimes 1)$, and $|\psi\rangle$, the state vector

$$|\psi\rangle = \left(1/\sqrt{2}\right)(1 \otimes 0) + \left(1/\sqrt{2}\right)(0 \otimes 1) = \begin{pmatrix} 0 \\ \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \\ 0 \end{pmatrix} .$$

The corresponding density reads

$$\rho = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} .$$

With a first measure being performed and corresponding to the operator A such that

$$A = \sigma_x^1 \otimes \sigma_y^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} ,$$

one gets the reduced density

$$\mathit{\Gamma}_A(\rho) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} .$$

Let the operator B be given as

$$B = \begin{pmatrix} \cos\frac{\pi}{4} & e^{-i\frac{\pi}{4}}\sin\frac{\pi}{4} \\ e^{i\frac{\pi}{4}}\sin\frac{\pi}{4} & -\cos\frac{\pi}{4} \end{pmatrix} \otimes \begin{pmatrix} \cos\frac{\pi}{3} & \sin\frac{\pi}{3} \\ \sin\frac{\pi}{3} & -\cos\frac{\pi}{3} \end{pmatrix} .$$

The eigenvalues of B are ± 1, the probability distribution of $X_B(\rho)$ is given by $p(X_B(\rho) = 1) \simeq 0.539\ldots$, whereas the probability distribution of $X_B(\mathit{\Gamma}_A(\rho))$ is given by $p(X_B(\mathit{\Gamma}_A(\rho)) = 1) \simeq 0.323\ldots$.

Now, if X is a rv such that $P(X = 1) = p$ and $P(X = -1) = 1-p$, then, $E(X) = 2p-1$, and $\mathrm{Var}(X) = E(X^2)-(E(X))^2 = 1-(2p-1)^2 = 4p(1-p)$. Therefore, one gets $\mathrm{Var}(X_B(\rho)) \simeq 0.993\ldots$, whereas $\mathrm{Var}(X_B(\mathit{\Gamma}_A(\rho))) \simeq 0.875\ldots$, showing that the dispersion on B has been somewhat reduced by the measure of A.

In spite of this reduction on the dispersion, we have not been able to work out an example where one would have

$$\sigma(X_A(\rho))\,\sigma(X_{B\leftarrow A}(\rho)) < \tfrac{1}{2}\left|\mathrm{Tr}(\rho\,(A\,B - B\,A))\right| .$$

7.3. On the symmetric inequality conservation

Contrarily to the standard (antisymmetric) Heisenberg inequality which, as we have just seen, does not survive the very first step of a sequential measurement process, it is worth noticing that the symmetrical component of the dispersion inequality remains

$$\sigma(X_A(\rho))\,\sigma(X_{B\leftarrow A}(\rho)) \geq \left|\tfrac{1}{2}\,\mathrm{Tr}(\mathit{\Gamma}_A(\rho)(A\,B+B\,A)) - \langle A\rangle_\rho\,\langle B\rangle_{\mathit{\Gamma}_A(\rho)}\right|$$

which, using (6), can equally be written as

$$\sigma(X_A(\mathit{\Gamma}_A(\rho)))\,\sigma(X_B(\mathit{\Gamma}_A(\rho))) \geq |\mathrm{Tr}(\mathit{\Gamma}_A(\rho)\frac{(AB+BA)}{2}) - \langle A\rangle_{\mathit{\Gamma}_A(\rho)}\,\langle B\rangle_{\mathit{\Gamma}_A(\rho)}\,|\,.$$

Up to the replacement of ρ by $\mathit{\Gamma}_A(\rho)$, the form of (23) is clearly preserved.

Of course, both right-hand sides need not be the same in the general case, though they can also be the same. For example, if A has eigenvalues $+1$ and -1, then $A = p_{A_{+1}} - p_{A_{-1}}$, $\mathbb{1} = p_{A_{+1}} + p_{A_{-1}}$, and $p_{A_{\pm 1}} = (\mathbb{1} \pm A)/2$. Thus the reduced density is

$$\mathit{\Gamma}_A(\rho) = \frac{(\mathbb{1}+A)}{2}\,\rho\,\frac{(\mathbb{1}+A)}{2} + \frac{(\mathbb{1}-A)}{2}\,\rho\,\frac{(\mathbb{1}-A)}{2} = \tfrac{1}{2}\,\rho + \tfrac{1}{2}\,A\,\rho\,A\,,$$

and then,

$$\begin{aligned}\tfrac{1}{2}\,\mathrm{Tr}(\mathit{\Gamma}_A(\rho)\,(AB+BA)) &= \tfrac{1}{4}\,\mathrm{Tr}(\rho(AB+BA)) + \tfrac{1}{4}\,\mathrm{Tr}(A\,\rho\,A(AB+BA)) \\ &= \tfrac{1}{2}\,\mathrm{Tr}(\rho(AB+BA))\,.\end{aligned}$$

That is

$$\sigma(X_A(\rho))\,\sigma(X_{B\leftarrow A}(\rho)) \geq \left|\tfrac{1}{2}\,\mathrm{Tr}(\rho\,(AB+BA)) - \langle A\rangle_\rho\,\langle B\rangle_{\rho_A}\right|$$

and, if moreover $\langle A\rangle_\rho = 0$, then, in this very situation, the symmetric inequality is conserved both in form and in magnitude.

8. On Bell's inequalities

8.1. Sequence of four measurements

As illustrated on the following array, a succession of measurements can be represented by the action of the intertwinning operator $\mathit{\Gamma}$:

$$\begin{array}{ccccc}
\text{observable} & \text{state} & \text{rv} & & \text{mean value} \\
 & \rho & & & \\
A & \downarrow & X_A(\rho) & \to & (\rho|A) \\
 & \mathit{\Gamma}_A(\rho) & & & \\
B & \downarrow & X_{B\leftarrow A}(\rho) = X_B(\mathit{\Gamma}_A(\rho)) & \to & (\rho|\mathit{\Gamma}_A B) \\
 & \mathit{\Gamma}_B\mathit{\Gamma}_A(\rho) & & & \\
C & \downarrow & X_{C\leftarrow B\leftarrow A}(\rho) = X_C(\mathit{\Gamma}_B\mathit{\Gamma}_A(\rho)) & \to & (\rho|\mathit{\Gamma}_A\mathit{\Gamma}_B C) \\
 & \mathit{\Gamma}_C\mathit{\Gamma}_B\mathit{\Gamma}_A(\rho) & & & \\
D & & X_{D\leftarrow C\leftarrow B\leftarrow A}(\rho) = X_D(\mathit{\Gamma}_C\mathit{\Gamma}_B\mathit{\Gamma}_A(\rho)) & \to & (\rho|\mathit{\Gamma}_A\mathit{\Gamma}_B\mathit{\Gamma}_C C)
\end{array}$$

In the sequel of Postulate 3, we now posit the following:

Postulate 4

The set of rvs $\{X_A(\rho), X_B(\Gamma_A(\rho)), X_C(\Gamma_B\Gamma_A(\rho)), X_D(\Gamma_C\Gamma_B\Gamma_A(\rho))\}$ admits the JPD given by

$$\begin{aligned} p(X_A(\rho)=a_i, X_B(\Gamma_A(\rho)) &= b_j, X_C(\Gamma_B\Gamma_A(\rho))=c_k, X_D(\Gamma_C\Gamma_B\Gamma_A(\rho))=d_l) \\ &= \mathrm{Tr}(P_{C_k}P_{B_j}P_{A_i}\,\rho\,P_{A_i}P_{B_j}P_{C_k}P_{D_l})\,. \end{aligned} \tag{28}$$

It is important to remark that this JPD allows one to obtain the JPDs of $(X_A(\rho), X_B(\Gamma_A(\rho)))$ and $(X_A(\rho), X_B(\Gamma_A(\rho)), X_C(\Gamma_B\Gamma_A(\rho)))$. This is so because,

$$\begin{aligned} & p(X_A(\rho) = a_i, X_B(\Gamma_A(\rho)) = b_j) \\ & = \sum_{k,l} p(X_A(\rho) = a_i, X_B(\Gamma_A(\rho)) = b_k, X_C(\Gamma_B\Gamma_A(\rho)) \\ & = c_k, X_D(\Gamma_C\Gamma_B\Gamma_A(\rho)) = d_l) = \sum_{k,l} \mathrm{Tr}(P_{C_k}P_{B_j}P_{A_i}\,\rho\,P_{A_i}P_{B_k}P_{C_k}P_{D_l}) \\ & = \mathrm{Tr}(P_{A_i}\,\rho\,P_{A_i}P_{B_j})\,, \end{aligned}$$

and

$$\begin{aligned} & p(X_A(\rho) = a_i, X_B(\Gamma_A(\rho)) = b_j, X_C(\Gamma_B\Gamma_A(\rho)) = c_k) \\ & = \sum_{l} p(X_A(\rho) = a_i, X_B(\Gamma_A(\rho)) = b_j, X_C(\Gamma_B\Gamma_A(\rho)) \\ & = c_k, X_D(\Gamma_C\Gamma_B\Gamma_A(\rho)) = d_l) = \sum_{l} \mathrm{Tr}(P_{C_k}P_{B_j}P_{A_i}\,\rho\,P_{A_i}P_{B_j}P_{C_k}P_{D_l}) \\ & = \mathrm{Tr}(P_{B_j}P_{A_i}\,\rho\,P_{A_i}P_{B_j}P_{C_k})\,. \end{aligned}$$

Remark Note that the situation which is envisaged here is at variance with the one of Fine's Theorem 3 [1], where, if there is a probability distribution on $\mathbb{R}^4$ that would be compatible with the probability distributions of rvs, $X_A(\rho)$, $X_{A'}(\rho)$, $X_B(\rho)$ and $X_{B'}(\rho)$ and all four probability distributions corresponding to the four couples $(X_A(\rho), X_B(\rho))$, $(X_A(\rho), X_{B'}(\rho))$, $(X_{A'}(\rho), X_B(\rho))$ and $(X_{A'}(\rho), X_{B'}(\rho))$, then we must have a set of inequalities, the Bell/CH inequalities. And in fact, there is equivalence between the existence of such a probability distribution and the Bell/CH inequalities. Note that in this case, the rvs, $(X_A(\rho), X_B(\rho))$, $(X_A(\rho), X_{B'}(\rho))$ can be represented by functions in a probability space, as in the Kolmogorov formalism of probability theory (this is a hidden variable theory in the sense of Ref. [1]).

Theorem 4 *Let four hermitian operators A, A', B and B' be such that the "As" commute only with the "Bs". Then, for the JPD of $X_A(\rho)$, $X_B(\Gamma_A(\rho))$, $X_C(\Gamma_B\Gamma_A(\rho))$ and $X_D(\Gamma_C\Gamma_B\Gamma_A(\rho))$, the following relations hold true*

$$
\begin{aligned}
E(X_A(\rho)X_{B\leftarrow A'\leftarrow A}(\rho)) &= \mathrm{Tr}(\rho\, A\, B)\,, \\
E(X_{A'\leftarrow A}(\rho)\ X_{B\leftarrow A'\leftarrow A}(\rho)) &= \mathrm{Tr}(\rho\, \Gamma_A(A')B)\,, \\
E(X_A(\rho)\ X_{B'\leftarrow B\leftarrow A'\leftarrow A}\,(\rho)) &= \mathrm{Tr}(\rho\, A\, \Gamma_B(B'))\,, \\
E(X_{A'\leftarrow A}(\rho)\, X_{B'\leftarrow B\leftarrow A'\leftarrow A}(\rho)) &= \mathrm{Tr}(\rho\, \Gamma_A(A')\, \Gamma_B(B'))\,.
\end{aligned}
$$

Proof: One will prove the first relation only, the three other's proof going the same way

$$
\begin{aligned}
E(X_A(\rho)\, X_{B\leftarrow A'\leftarrow A}(\rho)) &= \sum_{i,j,k} a_i\, b_k\, P_{B\leftarrow A'\leftarrow A,\rho}(a_i, {a'}_j, b_k) \\
&= \sum_{i,j,k} a_i\, b_k\, \mathrm{Tr}(P_{{A'}_j} P_{A_i}\, \rho\, P_{A_i} P_{{A'}_j} P_{B_k}) \\
&= \sum_{i,j,k} a_i\, b_k\, \mathrm{Tr}(P_{{A'}_j} P_{A_i}\, \rho\, P_{A_i} P_{B_k}\, P_{{A'}_j}) \\
&= \sum_{i,k} a_i\, b_k\, \mathrm{Tr} \sum_j P_{{A'}_j} (P_{A_i}\, \rho\, P_{A_i} P_{B_k}) \\
&= \mathrm{Tr}\left(\sum_i a_i\, (\rho\, P_{A_i}) \sum_k b_k P_{B_k} \right) \\
&= \mathrm{Tr}(\rho\, A\, B)\,.
\end{aligned}
$$

End of proof.

Theorem 4 will be helpful in order to analyze a Bell's type inequality in a sequential context. Not surprisingly, we will see then, that results come out fairly different from what is ordinarily put forth and that is briefly reminded in the next subsection.

8.2. A standard Bell inequality and its quantum violation

Let us consider the case of four bivalent (± 1) random variables: X, X', Y and Y', and *assume that we have a JPD for* (X, X', Y, Y'). Thus we rely on the hypothesis that, for each instance of the experiments, the numbers X, X', Y, Y' are determined. This means that for each instance, a combination like $Z = (X + X')\, Y + (X - X')\, Y'$ is well defined. By inspection of all the possible values of X, X', Y and Y' one has

$$
\begin{aligned}
(X + X') &= \pm\, 2 \Longleftrightarrow (X - X') = 0\,, \\
(X + X') &= 0 \Longleftrightarrow (X - X') = \pm\, 2\,,
\end{aligned}
$$

and so, $|Z| \leq 2$. Taking the mean value, the following *Bell inequality* results trivially

$$|E(Z)| = |E(XY) + E(X'Y) + E(XY') - E(X'Y')| \leq 2\,. \tag{29}$$

We remark that this inequality does not depend of the explicit form of the JPD of the four random variables.

Consider now a quantum mechanical situation where one has:

— Four observables A, A', B and B', all four having ± 1 as their eigenvalues,

— The As commute with the Bs, but $[A, A'] \neq 0$ and $[B, B'] \neq 0$.

Because of the commutation of As with the Bs, in any state ρ, JPDs do exist for the 4 couples of rvs, $(X_A(\rho), X_B(\rho))$, $(X_A(\rho), X_{B'}(\rho))$, $(X_{A'}(\rho), X_B(\rho))$ and $(X_{A'}(\rho), X_{B'}(\rho))$. But because $[A, A'] \neq 0$ and $[B, B'] \neq 0$, we do not have a JPD for $(X_A(\rho), X_B(\rho), (X_{A'}(\rho), X_{B'}(\rho))$, in view of Fine's Theorem 7, [1].

As a matter of fact, a combination like

$$Z = (X_A(\rho) + X_{A'}(\rho))X_B(\rho) + (X_A(\rho) - X_{A'}(\rho))X_{B'}(\rho)\,,$$

does not define any rv, even though the 4 products $X_A(\rho)\, X_B(\rho)$, $X_{A'}(\rho)$ $X_B(\rho)$, $X_{A'}(\rho)\, X_B(\rho)$ and $X_{A'}(\rho)\, X_{B'}(\rho)$ do define 4 rvs, separately. This means that an expression like

$$|E(X_A(\rho)X_B(\rho)) + E(X_{A'}(\rho)X_B(\rho)) + E(X_A(\rho)X_{B'}(\rho)) - E(X_{A'}(\rho)X_{B'}(\rho))|\,,$$

whose terms are meaningful, does not necessarily satisfies (29).

And effectively, there exist particular states, dubbed *entangled*, where the absolute value written above is found to display values greater than 2 (up to $2\sqrt{2}$, in that very case). It is this amazing result of Quantum Mechanics that has been verified experimentally [11], and in the context of a finite speed limit, has given rise to the notion of *failure of the local realism hypothesis* and to the notion of *quantum non-separability*.

8.3. Bell's inequalities versus sequential measurements

Let four observables A, A', B and B', have eigenvalues $+1$ and -1, not all of them commuting two by two: $[A\text{s}, B\text{s}] = 0$, $[A, A'] \neq 0$, $[B, B'] \neq 0$. If measurements are performed sequentially, then by Postulate 4 and by (29), the following inequality holds necessarily,

$$\begin{aligned}&|E(X_A(\rho)\; X_{B\leftarrow A'\leftarrow A}(\rho) + X_{A'\leftarrow A}(\rho)\; X_{B\leftarrow A'\leftarrow A}(\rho) \\ &+X_A(\rho)\; X_{B'\leftarrow B\leftarrow A'\leftarrow A}(\rho) - X_{A'\leftarrow A}(\rho)\; X_{B'\leftarrow B\leftarrow A'\leftarrow A}(\rho))| \leq 2\,.\end{aligned} \tag{30}$$

As remarked after (29), the specific form of the four rvs JPD involved in (30) is not required for the latter inequality to be satisfied. However, by Postulate 4, we do have that very JPD. For the sake of consistency it is therefore a crucial step to check that a calculation based on the explicit JPD of Postulate 4, complies with (30). That this consistency holds is the matter of the following two results.

Theorem 5 *Let four hermitian operators, A, A', B and B', be such that the "As" commute only with the "Bs". Then*

$$\begin{aligned}&E(X_A(\rho)\ X_{B\leftarrow A'\leftarrow A}(\rho) + X_{A'\leftarrow A}(\rho)\ X_{B\leftarrow A'\leftarrow A}(\rho)\\ &+X_A(\rho)\ X_{B'\leftarrow B\leftarrow A'\leftarrow A}(\rho) - X_{A'\leftarrow A}(\rho)\ X_{B'\leftarrow B\leftarrow A'\leftarrow A}(\rho)) = \mathrm{Tr}(\rho\, K)\,,\end{aligned} \tag{31}$$

where

$$K = A(B + \Gamma_B(B')) + \Gamma_A(A')(B - \Gamma_B(B'))\,.$$

This is a mere consequence of previous Theorem 4. Then, in a second place

Theorem 6 *Theorem of consistency. If $K = A(B+\Gamma_B(B')) + \Gamma_A(A')(B - \Gamma_B(B'))$, for hermitian operators, A, A', B, B', having ± 1 as their eigenvalues and such that the* As *commute solely with the* Bs, *then, for any statistical operator ρ, one has*

$$|\mathrm{Tr}(\rho\, K)| \leq 2\,.$$

Before giving the proof for this theorem, let us begin with recalling that for bounded linear operators A on a vectorial space $\mathbb{E}$, we have the norm $\|\ \|_{\mathcal{L}}$

$$\|A\|_{\mathcal{L}} = \sup_{\|\psi\|=1} \|A\,\psi\| = \sup_{\psi\in\mathbb{E}} \frac{\|A\psi\|}{\|\psi\|}\,.$$

If $\mathbb{E}$ is a $\mathbb{C}$–Hilbert space, then, in virtue of the Cauchy–Schwartz inequality, $|\langle\phi|A\psi\rangle| \leq \|A\psi\| \times \|\phi\|$, an equivalent form of the norm is given by

$$\|A\|_{\mathcal{L}} = \sup_{\|\psi\|=1\,,\ \|\phi\|=1} |\langle\phi|A\psi\rangle|\,.$$

For any two linear bounded operators A and B on $\mathbb{E}$, the following two inequalities hold

$$\|A+B\|_{\mathcal{L}} \ \leq\ \|A\|_{\mathcal{L}} + \|B\|_{\mathcal{L}}\,, \tag{32}$$

$$\|A\,B\|_{\mathcal{L}} \ \leq\ \|A\|_{\mathcal{L}} \times \|B\|_{\mathcal{L}}\,, \tag{33}$$

where of a first lemma.

Lemma 2 *Let ρ be a statistical operator, and K a continuous linear operator on a $\mathbb{C}$–Hilbert space $\mathbb{E}$, then*

$$|\mathrm{Tr}(\rho\, K)| \leq \|K\|_{\mathcal{L}}\,.$$

Proof: Being hermitian, ρ can be diagonalized in a basis of orthonormal eigenvectors, $\{\psi_i\}$, with positive eigenvalues, $\rho_i \equiv \langle\psi_i|\rho\,\psi_i\rangle$. We denote $K_i \equiv \langle\psi_i|K\,\psi_i\rangle$. Then,

$$|\mathrm{Tr}(\rho\, K)| = \left|\sum_{i=1}^{n} \rho_i\, K_i\right| \leq \sum_{i=1}^{n} \rho_i\, |K_i|\,.$$

Now, $|K_i| \leq \|K\|_{\mathcal{L}}$ gives

$$|\mathrm{Tr}(\rho\, K)| \leq \|K\| \left(\sum_{i=1}^{n} \rho_i\right) = \|K\|_{\mathcal{L}}\,,$$

since $\mathrm{Tr}(\rho) = \sum\limits_{i=1}^{n} \rho_i = 1$. End of proof.

And another Lemma.

Lemma 3 *Let A be a linear continuous hermitian operator, and B a continuous linear operator on a $\mathbb{C}$–Hilbert space $\mathbb{E}$. Writing $A \equiv \sum\limits_k a_k\, P_k$ at $\sum\limits_k P_k = \mathbb{1}$, where the $a_k \in \mathbb{R}$ are the eigenvalues of A, and P_k the projector on the subspace associated to a_k, then*

$$\|\mathit{\Gamma}_A(B)\|_{\mathcal{L}} \leq \|B\|_{\mathcal{L}}\,.$$

If $B \equiv \rho$ is a density matrix, then the equalities $\|\mathit{\Gamma}_A(\rho)\|_{\mathcal{L}} = \|\rho\|_{\mathcal{L}} = 1$ hold.

Proof: Let $\psi \in \mathbb{E}$. We write $\psi_k = P_k\,\psi$. One has

$$\|\mathit{\Gamma}_A(B)\psi\|^2 = \left\|\sum_k (P_k B P_k)\psi_k\right\|^2 = \sum_k \left\|(P_k B P_k)\psi_k\right\|^2.$$

From (33), one has $\|P_k B P_k\|_{\mathcal{L}} \leq \|B\|_{\mathcal{L}}$, since for all k, $\|P_k\|_{\mathcal{L}} = 1$. And so,

$$\|\mathit{\Gamma}_A(B)\psi\|^2 \leq \|B\|^2_{\mathcal{L}} \left(\sum_k \|\psi_k\|^2\right) = \|B\|^2_{\mathcal{L}} \,.\, \|\psi\|^2\,.$$

End of proof.

This allows to derive easily the result.

Lemma 4 *Let A, A', B and B' be 4 continuous linear hermitian operators on a $\mathbb{C}$–Hilbert space $\mathbb{E}$, with eigenvalues ±1, and such that the As commute with the Bs; and let K be the operator*

$$K = A(B + \Gamma_B(B')) + \Gamma_A(A')(B - \Gamma_B(B'))\,,$$

then $\|K\|_{\mathcal{L}} \leq 2$.

Proof of Theorem 6: Let $\psi \in \mathbb{E}$, $A = \sum a_i P_{A_i}$ and $B = \sum b_k P_{B_k}$, where P_{A_i} and P_{B_k} are the projectors on the eigensubspaces corresponding respectively to the eigenvalues $a_i = \pm1$ and $b_k = \pm1$, with i, $k = \pm$. They satisfy to the relations $\sum\limits_i P_{A_i} = \mathbb{1}$ and $\sum\limits_k P_{B_k} = \mathbb{1}$. The operator K can be written as

$$K = \sum_{i,k\,=\,\pm} Q_{i,k} K_{i,k} Q_{i,k}\,,$$

where $Q_{i,k} = P_{A_i} P_{B_k}$, and where

$$K_{i,k} = a_i(b_k\,\mathbb{1} + P_{B_k} B' P_{B_k}) + P_{A_i} A' P_{A_i}(b_k\,\mathbb{1} - P_{B_k} B' P_{B_k})\,.$$

One has $\sum\limits_{i,k\,=\,\pm} Q_{i,k} = \mathbb{1}$, $Q_{i,k}^2 = Q_{i,k}$ and $Q_{i,k}\,Q_{i',k'} = \delta_{ii'}\delta_{kk'}$, showing that the $Q_{i,k}s$ are complementary projectors on supplementary eigensubspaces spanning $\mathbb{E}$. In particular, since for all i and k, $\|Q_{i,k}\|_{\mathcal{L}} = 1$, one gets

$$\|K\psi\|^2 = \sum_{i,k\,=\,\pm} \|Q_{i,k} K_{i,k} Q_{i,k}\,\psi\|^2 \leq \sum_{i,k\,=\,\pm} \|K_{i,k}\|_{\mathcal{L}}^2\,\|\psi_{i,k}\|^2\,, \tag{34}$$

where the $\psi_{i,k} \equiv Q_{i,k}\psi$ are orthogonal. Thus, $\sum\limits_{i,k\,=\,\pm} \psi_{i,k} = \psi$, and so, $\sum\limits_{i,k\,=\,\pm} \|\psi_{i,k}\|^2 = \|\psi\|^2$. Now, one has $\|K_{i,k}\|_{\mathcal{L}} \leq 2$. This can be seen as follows $P_{B_k} B' P_{B_k}$ and $P_{A_i} A' P_{A_i}$ commute. Therefore, an orthonormal basis of common eigenvectors can be found. Let $\phi \in \mathbb{E}$. This vector can be written as $\phi = \sum\limits_s \varphi_{(i,k),s}$, where the $\varphi_{(i,k),s}$ are the orthogonal eigenvectors of both $P_{B_k} B' P_{B_k}$ and $P_{A_i} A' P_{A_i}$, with respective eigenvalues $p_{(i,k),s}$ and $q_{(i,k),s}$. One has $|p_{(i,k),s}| \leq 1$ and likewise $|q_{(i,k),s}| \leq 1$ since $\|P_{B_k} B' P_{B_k}\|_{\mathcal{L}} \leq \|B'\|_{\mathcal{L}} = 1$ and $\|P_{A_i} A' P_{A_i}\|_{\mathcal{L}} \leq \|A'\|_{\mathcal{L}} = 1$.

One can write,

$$K_{i,k}\,\phi = \sum_s \Big(a_i(b_k + p_{(i,k),s}) + q_{(i,k),s}(b_k - p_{(i,k),s})\Big)\varphi_{(i,k),s}\,,$$

and denote $\alpha_{(i,k),s}$, the quantity

$$\alpha_{(i,k),s} \equiv a_i(b_k + p_{(i,k),s}) + q_{(i,k),s}(b_k - p_{(i,k),s})\,.$$

At $a_i = \pm 1$ and $b_k = \pm 1$, one can write

$$|\alpha_{(i,k),s}| \leq |a_i|\,|b_k + p_{(i,k),s}| + |q_{(i,k),s}|\,|b_k - p_{(i,k),s}| \leq |b_k + p_{(i,k),s}| + |b_k - p_{(i,k),s}|\,.$$

Now, if $b_k = +1$, then, $|b_k + p_{(i,k),s}| + |b_k - p_{(i,k),s}| = |1 + p_{(i,k),s}| + |1 - p_{(i,k),s}|$, and if $b_k = -1$, then, $|b_k + p_{(i,k),s}| + |b_k - p_{(i,k),s}| = |-1 + p_{(i,k),s}| + |-1 - p_{(i,k),s}| = |1 - p_{(i,k),s}| + |1 + p_{(i,k),s}|$. Since $|p_{(i,k),s}| \leq 1$, then, $|1 + p_{(i,k),s}| + |1 - p_{(i,k),s}| = 1 + p_{(i,k),s} + 1 - p_{(i,k),s} = 2$, and thus, for any b_k-value, one has $|\alpha_{(i,k),s}| \leq 2$, and accordingly

$$\|K_{i,k}\,\phi\|^2 = \sum_s |\alpha_{(i,k),s}|^2\,\|\varphi_{(i,k),s}\|^2 \leq \sum_s 4\,\|\varphi_{(i,k),s}\|^2 = 4\|\phi\|^2\,.$$

That is, $\|K_{i,k}\,\phi\| \leq 2\|\phi\|$, or $\|K_{i,k}\|_{\mathcal{L}} \leq 2$. Eventually (34) can be put into the form

$$\|K\psi\|^2 \leq 4 \sum_{i,k=\pm} \|\psi_{i,k}\|^2 = 4\,\|\psi\|^2\,,$$

that is, $\|K\psi\| \leq 2\|\psi\|$, or $\|K\|_{\mathcal{L}} \leq 2$. End of proof.

A fundamental example In order to illustrate these somewhat formal consistency results with an example, one may consider the particular case of 4 spin operators given by

$$A = \sigma_z \otimes Id\,, \qquad A' = \sigma_x \otimes Id\,,$$
$$B = -\frac{1}{\sqrt{2}}\,Id \otimes (\sigma_z + \sigma_x)\,, \qquad B' = \frac{1}{\sqrt{2}}\,Id \otimes (\sigma_z - \sigma_x)\,,$$

where $[A, A'] \neq 0$ and $[B, B'] \neq 0$, whereas the A's commute with the B's. In order to emphasize the difference with the famous case of Bell's inequalities violation, we can choose the entangled spin zero singlet state which reads

$$\rho_s = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Now, with X_Z the rv,

$$X_Z = X_A(\rho_s)\;X_{B\leftarrow A'\leftarrow A}(\rho_s) \;+\; X_{A'\leftarrow A}\,(\rho_s)\,X_{B\leftarrow A'\leftarrow A}(\rho_s)$$
$$+\; X_A\,(\rho_s)\,X_{B'\leftarrow B\leftarrow A'\leftarrow A}\,(\rho_s) \;-\; X_{A'\leftarrow A}\,(\rho_s)\,X_{B'\leftarrow B\leftarrow A'\leftarrow A}\,(\rho_s)\,,$$

one has $E(X_Z) = \mathrm{Tr}(\rho_s\,K)$ in view of Theorem 5, and an explicit calculation yields in effect

$$E(X_Z) = \frac{\sqrt{2}}{2}\,.$$

Thus,

$$
\begin{aligned}
&|E(X_A(\rho_s)\ X_{B\leftarrow A'\leftarrow A}(\rho_s)) + E(\ X_{A'\leftarrow A}\,(\rho_s)\,X_{B\leftarrow A'\leftarrow A}(\rho_s)) \\
&+E(\ X_A\,(\rho_s)\,X_{B'\leftarrow B\leftarrow A'\leftarrow A}\,(\rho_s)) \\
&-E(\ X_{A'\leftarrow A}\,(\rho_s)\,X_{B'\leftarrow B\leftarrow A'\leftarrow A}\,(\rho_s))| < 2\,,
\end{aligned}
\tag{35}
$$

as expected in view of Theorem 6, whereas it is well known that in the (maximally) entangled state ρ_s one has instead,

$$
\begin{aligned}
&|E(X_A(\rho_s)\ X_B(\rho_s)) + E(\ X_{A'}(\rho_s)\ X_B(\rho_s)\,) \\
&+E(\ X_A\,(\rho_s)\,X_{B'}(\rho_s)) - E(\ X_{A'}(\rho_s)\ X_{B'}(\rho_s))| = 2\sqrt{2}\,.
\end{aligned}
\tag{36}
$$

9. Conclusion

In this article, two among the most famous inequalities of Quantum Mechanics, that are Heisenberg and Bell inequalities, have been considered in the particular context of sequential measurements. Apart from an intrinsic theoretical (if not experimental, [10]) interest on its own, these considerations are also motivated by the long recognized *contextual* aspect of Quantum Mechanics [12], which needs to be explored so as to grasp the largest overview of the *quantum specificity*.

The Heisenberg inequality splits naturally into its real (symmetric) and imaginary (antisymmetric) parts. Along a sequential measurement process the antisymmetric form gets a right-hand side which is identically zero at the very first step of the process, whereas the symmetric part of it remains unaltered in form. Moreover, even though rvs dispersions would seem to be reduced through a sequential procedure, we have not been able to see whether the Heisenberg inequality lower boundary can itself be somewhat reduced in the course of such a process.

We have seen here also the substantial change brought about by the sequential measurement context, since JPDs can be defined, and Bell's inequalities satisfied accordingly, even though one starts from entangled states (even *maximally*), and considers the case of operators that do not all commute two by two. A peculiar aspect of the sequential measurement procedure is therefore that it provides sequential (quantum) random variables with Joint Probability Distributions, even though not all of the corresponding quantum operators commute. In other words, in a sequential measurement context, quantum inequalities become identical to classical ones.

Throughout the article we have been relying on notations that may have appeared a bit queer if not cumbersome sometimes: The $X_{B\leftarrow A\leftarrow\ldots}(\rho)$, corresponding to (quantum) peculiar random variables. These notations, though, have merits which deserve to be emphasized:

— In the first place, they operate a clear-cut and necessary distinction between observables, represented by hermitian operators, A, (possibly enlarged to so-called *pseudo-hermitian* operators [13]) on some Hilbert space of states, $\mathbb{E}$, and $X_A(\rho)$, the associated random variable, depending on the state ρ.

— So defined, these (quantum) random variables should not be mistaken for the Kolmogorov formalism homonymous variables, adapted to situations where the local realism hypothesis is not questioned.

— As compared to $X_B(\rho)$, the notation $X_{B\leftarrow A}(\rho)$ allows to express a (quantum) random variable of a different type since it is not associated to any new operator than B, but to B within a process, and may be dubbed a *sequential random variable*. Not only does this account for the contextual aspect of Quantum Mechanics, but also, used properly, it helps dissipating easily some apparent paradoxical aspects, such as those of the *Three boxes paradox* for example [14, 15]. In order to alleviate the text, this example and related extensions of the present analysis, have not been dealt with in the present article and could be presented elsewhere.

We wish to thank the unknown Referee whose criticisms have helped us to improve a previous version of the present article.

REFERENCES

[1] A. Fine, *J. Math. Phys.* **23**, 1306 (1982).

[2] A.M. Gleason, *J. Math. Mech.* **6**, 885 (1957).

[3] W.H. Zurek, *Rev. Mod. Phys.* **75**, 715 (2003).

[4] M.Le Bellac, *Physique Quantique*, CNRS, ed. EDP Sciences, Paris 2007.

[5] E.P. Wigner, *Am. J. Phys* **31**, 6 (1963).

[6] R.B. Griffiths, *J. Statist. Phys.* **36**, 219 (1984).

[7] M.D. Srinivas, *Pramana* **60**, 1137 (2003).

[8] A. Khrennikov, *Information Science* **179**, 492 (2009) [`arXiv:0709.3909v2 [quant-ph]`].

[9] B. Gnanapragasam, M.D. Srinivas, *Pramana* **12**, 669 (1979).

[10] N. Katz *et al.*, *Phys. Rev. Lett* **101**, 200401 (2008) [`arXiv:0806.3547v1 [quant-ph]`].

[11] A. Aspect, J. Dalibard, G. Roger, *Phys. Rev. Lett.* **49**, 1804 (1982).

[12] M. Bitbol, *Mécanique quantique: Une introduction philosophique*, Flammarion, Paris, France, 1996.

[13] C.M. Bender, S.P. Klevansky, *Phys. Lett.* **A373**, 2670 (2009).

[14] Y. Aharonov, L. Vaidman, *Phys. Rev.* **A41**, 11 (1990).

[15] L. Vaidman, *Found. Phys.* **26**, 895 (1996).