

Born Rule and Noncontextual Probability

Fabrizio Logiurato^{1,2*}, Augusto Smerzi²

¹INO-CNR BEC Center and Physics Department, Trento University, Povo, Italy

²QSTAR, INO-CNR and LENS, Largo Enrico Fermi, Firenze, Italy

Email: *log@science.unitn.it

Received September 2, 2012; revised October 2, 2012; accepted October 10, 2012

ABSTRACT

We present a new derivation of the Born rule from the assumption of noncontextual probability (NCP). Within the theorem we also demonstrate the continuity of probability with respect to the amplitudes, which has been suggested to be a gap in Zurek's and Deutsch's approaches, and we show that NCP is implicitly postulated also in their derivations. Finally, physical motivations of NCP are given based on an invariance principle with respect to a resolution change of measurements and with respect to the principle of no-faster-than-light signalling.

Keywords: Born Rule; Non Contextual Probability

1. Introduction

The fundamental probabilistic postulate which allows QM its successful predictions was first introduced by Born (from a suggestion by Einstein) in 1926 within the scattering theory [1]. The question if such postulate could be derived as a theorem from other postulates of QM was quickly posed. A first answer was provided by Von Neumann and included in his famous book of 1932 on the mathematical foundations of QM [2]. His theorem nowadays does not enjoy much attention since it is believed to contain more assumptions than Gleason's [3,4]. Different attempts have been discussed in the context of the relative-state [5] and the many-worlds interpretations [6] of QM by Finkelstein [7] and Hartle [8], which have analyzed an endless sequence of measurements to show that the relative frequency follows the Born rule. However, the meaning in the real world of an infinite sequence of measurements is controversial and it is doubtful whether these proofs contain circularities [9-14].

The 1957 Gleason work [4] is regarded as the most important derivation of the Born rule [15], but this theorem is usually considered quite formal and difficult to grasp. Moreover, it is based on the concept of noncontextual probability (NCP), whose connection with physics is not clear. In the structure of the Gleason theorem, every Hilbert subspace \mathcal{M} corresponds to an observable quantity. Each one of such subspaces is representable by a projection operator $\hat{\Pi}_{\mathcal{M}}$. A projector $\hat{\Pi} = |\psi\rangle\langle\psi|$ can represent the question with answer yes or not concerning an experiment testing whether the system has the respec-

tive state $|\psi\rangle$. A complete set of mutually commuting projectors $\{\hat{\Pi}_i\}$ corresponds to a set of questions which can simultaneously be asked in a measurement. In the logical-algebraic approach [16], probability is a measure defined over a projection lattice (set of closed subspaces) of Hilbert space \mathcal{H} , a mapping $p: \hat{\Pi}_{\mathcal{M}} \rightarrow [0,1]$. The postulates on which the theorem rests are:

- 1) The probability assigned to a complete set of N projectors $\{\hat{\Pi}_i\}$ is normalized:

$$p(I) = 1, \quad \sum_{i=1}^N \hat{\Pi}_i = I. \quad (1)$$

- 2) For any sequence of M mutually orthogonal projectors:

$$p\left(\sum_{i=1}^M \hat{\Pi}_i\right) = \sum_{i=1}^M p(\hat{\Pi}_i). \quad (2)$$

Assuming Postulate 2) is equivalent to postulating NCP for projective measurements (or, *i.e.*, according to the definition by Spekkens, measurement noncontextuality [17]): the probability of a certain occurrence, for example the answer yes for a projector, does not depend on other questions simultaneously tested, that is, on other projectors. In terms of basis vectors, the probability of obtaining a given state is independent of the basis it belongs to. To illustrate this important point, let us consider, for instance, two complete set of mutually orthogonal projectors $\{\hat{\Pi}_i\}$, $\{\hat{\Pi}'_i\}$ with $N = 3$, where $\hat{\Pi}_1 \equiv \hat{\Pi}'_1$. From the normalization postulate of probability:

$$p(\hat{\Pi}_1) + p(\hat{\Pi}_2) + p(\hat{\Pi}_3) = 1, \quad (3)$$

$$p'(\hat{\Pi}_1) + p(\hat{\Pi}'_2) + p(\hat{\Pi}'_3) = 1, \quad (4)$$

*Corresponding author.

where in principle $p(\hat{\Pi}_1)$ and $p'(\hat{\Pi}_1)$ may be different. In the subspace orthogonal to $\hat{\Pi}_1$ we have $\hat{\Pi}_2 + \hat{\Pi}_3 = \hat{\Pi}'_2 + \hat{\Pi}'_3$, then $p(\hat{\Pi}_2 + \hat{\Pi}_3) = p(\hat{\Pi}'_2 + \hat{\Pi}'_3)$ and from Postulate 2:

$$p(\hat{\Pi}_2) + p(\hat{\Pi}_3) = p(\hat{\Pi}'_2) + p(\hat{\Pi}'_3). \tag{5}$$

Then, by comparing Equation (3) and Equation (4), we get $p(\hat{\Pi}_1) = p'(\hat{\Pi}_1)$. It is also possible to show the vice versa, *i.e.*, if NCP holds then Postulate 2 is true.

From Postulates 1 and 2 it follows *Gleason's Theorem*: For a Hilbert space with dimension $\dim\mathcal{H} \geq 3$ there exists a density operator $\hat{\rho}$ which acts on \mathcal{H} such that the probability measure for every projector $\hat{\Pi}$ is the trace rule:

$$p(\hat{\Pi}) = \text{tr}(\hat{\rho}\hat{\Pi}), \tag{6}$$

namely, the Born rule.

More recently, new proofs appeared in the literature with the intention of being "more physically motivated than the theorem of Gleason" [18]. Their aim, as the ours in the present work, is to shed new light on the physical principles of QM, particularly on the origin of the Born rule. Indeed, Gleason's theorem gives "rather little insight into the emergence of quantum probabilities and the Born rule" [19].

Particularly interesting are Deutsch's and Zurek's derivations [18,20-25]. Their approaches to the Born rule have similar structures. As a first step, they use an invariance principle in order to show that equal amplitudes correspond to equal probabilities. Deutsch finds such principle in the theory of decisions (for critical discussions and further elaborations see [26-28]). Zurek introduces an environment-assisted invariance principle [29,30], *envariance*, in the framework of a relative-state approach where the system \mathcal{S} under measurement jointly evolves with the environment \mathcal{E} and, possibly, with one or more auxiliary measurement devices. After demonstrating the correspondence between equal amplitudes and equal probabilities, both Zurek and Deutsch consider a fine-graining technique to deal with the general case of different amplitudes of the initial state. For this purpose they have to introduce auxiliary systems which become entangled with \mathcal{S} .

However, fine-graining has been questioned: Caves [23] considers disturbing the necessity of introducing additional systems of adequate dimensions, possibly in an infinite Hilbert space, in order to reach the wanted approximation to irrational probabilities. Barnum [25] points out an even more stringent weakness: the step from rational to irrational amplitudes requires the continuity of the probability with respect to such amplitudes and such

a property is not demonstrated by Zurek and Deutsch. We note that a fundamental and extensive part of Gleason's theorem is actually devoted to prove that the probability is continuous. For several other controversial points of Zurek's demonstration, the most of them shares with Deutsch's, we refer to [19,23-25].

So, although Gleason-type theorems are usually considered well established, a simple, physically justified and shared derivation of the Born rule is not yet present in the literature. With our work we intend to fill this gap.

In the next section we propose a new derivation of the Born rule from the NCP assumption for projective measurements and pure states, the generalization to mixed states being immediate. Whereas Gleason assumes that a density operator acting on \mathcal{H} describes a general quantum state of the system only at the final step of his work, we introduce vectors of the Hilbert space describing quantum states from the very beginning of our derivation. On the one hand it allows a much more elementary deduction of the Born rule, and on the other hand it makes the postulates and the physical principles of our theorem easier to analyze in comparison with the previous demonstrations. Moreover, involving Zurek's and Deutsch's demonstrations pure states, also the connections between those works, the Gleason Theorem and the ours are clearer.

We do not enter into the controversy about the origin and the meaning of probability, and following Gleason and others (see also Mohrhoff in [24]), we assume the existence of probability defined on a quantum state. In addition, differently by Zurek, we do not suppose unitary evolution and consider a single system. As Gleason's Theorem, our result holds for a Hilbert space with $\dim\mathcal{H} \geq 3$, whereas in general, because of fine-graining, Zurek's and Deutsch's hold for $\dim\mathcal{H} = \infty$. We also give a rather simple derivation that the probability is continuous from NCP. Without this proof, Deutsch's and Zurek's results only hold for rational values of amplitudes.

In Section III we generalize our theorem to include entangled states and multi-particles systems. Since the Hilbert space \mathcal{H} of such systems has dimension which is the product of the dimensions of its components, the coupling of a subsystem with $\dim\mathcal{H} \geq 3$ with another subsystem with $\dim\mathcal{H} = 2$ enables us to extend the Born rule also to such space.

In Section IV we show that the Gleason postulate of NCP occurs in the use of fine-graining and that such a postulate is a hidden assumption of Zurek's and Deutsch's proofs. Indeed, fine-graining is a resolution change of our measurements and NCP is equivalent to a condition of invariance with respect to such a change: fine-graining on a subspace of the state must not alter the probability of finding the system in a different subspace.

This resolution invariance is an intuitive and physically natural interpretation of measurement noncontextuality.

Finally, in Section V, we show that the condition of noncontextual probability, or resolution invariance, can also be deduced by the hypothesis of no-faster-than-light signalling of the Special Relativity.

2. Born Rule as a Theorem from Noncontextual Probability

Contrarily to the principles of classical mechanics, the axioms of QM do not have an equally shared and clear formulation. We do not want to enter into such a thorny issue, related to the interpretative problems of QM. For our purpose, from the quantum postulates generally considered in standard presentations of QM (see, e.g., [31-34]), we can extract the following common assumptions:

Postulate I:

At the time t the state of every physical system \mathcal{S} is described by a normalized vector $|\psi_S(t)\rangle$ belonging to a Hilbert space \mathcal{H} .

Postulate II:

Every measurable physical quantity \mathcal{A} is described by a Hermitian operator \hat{A} acting in \mathcal{H} .

Postulate III:

The only possible result of a measurement of an observable \mathcal{A} is one of the eigenvalues $\{a_i\}$ of \hat{A} , (as the customary habit, from now on we will identify the observable with its operator).

Postulate IV:

If the measurement of the observable \hat{A} on the system in the state $|\psi_S(t)\rangle$ gives the result a_i , the state of the system immediately after the measurement is the associate normalized eigenstate $|a_i\rangle$ (non-degenerate case).

Postulate I is known as the *Completeness Postulate of QM*. *Postulate II* links an observable quantity with an operator in a Hilbert space (*observable-operator link*). *Postulate III* connects a particular value of that observable with an eigenvalue of the corresponding operator (*value-eigenvalue link*). *Postulate IV* relates a detected eigenvalue with the state of the system (*eigenvalue-eigenstate link*).

Now consider a Hermitian observable \hat{A} with discrete eigenvalues $\{a_i\}$ and eigenstates $\{|a_i\rangle\}$, $i=1, \dots, N$, and a single system \mathcal{S} described by the state $|\psi_S\rangle \in \mathcal{H}^N$. $\{|a_i\rangle\}$ is a complete orthonormal set and $|\psi_S\rangle$ can be decomposed into components:

$$|\psi_S\rangle = \sum_{i=1}^N c_i |a_i\rangle. \quad (7)$$

We intend to prove the following

Theorem 1:

If NCP and *Postulates I-IV* hold, then for $N \geq 3$ the

probability of obtaining the non-degenerate eigenvalue a_i , measuring \hat{A} on the state $|\psi_S\rangle$, is necessarily given by the square modulus of the inner product:

$$p(a_i) = |\langle a_i | \psi_S(t) \rangle|^2 \quad \text{Born Rule,} \quad (8)$$

where $|a_i\rangle$ is the eigenstate associated with the eigenvalue a_i .

We assume the probability as a primitive concept and we begin the derivation of *Theorem 1* by proving some simple lemmas. We underline that any theorem having the aim to prove the Born rule has to face the fact that the connection between probability and quantum state must be postulated or deduced by other means. Gleason, for instance, postulates the probability to be a measure defined on states of the basis. As Gleason, we suppose QM to be a probabilistic theory, which makes predictions about probabilities of occurrences of observable quantities, and that probabilities are functions of quantum states. Then, from QM postulates we want to show that

Lemma 1:

The probability of obtaining an eigenvalue a_i is equal to the probability of obtaining the respective eigenstate $|a_i\rangle$ (statistical case):

$$p(a_i) = p(|a_i\rangle). \quad (9)$$

Proof:

Suppose that the system is prepared with the eigenvalue a_i , in this case $p(a_i)=1$. According to *Postulate IV* if $p(a_i)=1$ then $p(|a_i\rangle)=1$, and vice versa. But $\{a_i\}$ correspond to mutually exclusive events, thus has to be $p(a_j)=0$ for $j \neq i$. Then must necessarily be $p(|a_j\rangle)=0$ too, since, otherwise, having obtained with a measurement the state $|a_j\rangle$ we should assign to the system the eigenvalue a_j , in contradiction with the exclusivity property of $\{a_i\}$ events. Therefore we have the equations for the conditional probabilities for a non-statistical case:

$$p(|a_i\rangle; a_j) = p(a_j; |a_i\rangle) = \delta_{ij}, \quad (10)$$

which represent the probability of finding the state $|a_i\rangle$ for systems prepared with certainty with eigenvalue a_j , and vice versa. Because of *Postulate III*, the only possible result of a measurement is one of the eigenvalues of \hat{A} . Consider then a generic preparation $|\psi_S\rangle$ of the initial state with $p(|a_i\rangle) \neq 1$. Introducing the joint probability $p(|a_i\rangle, a_j)$ and considering the general validity of the Bayesian formula for the joint probabilities $p(|a_i\rangle, a_j) = p(|a_i\rangle; a_j) p(a_j)$, we have

$$\begin{aligned} p(|a_i\rangle) &= \sum_{j=1}^N p(|a_i\rangle, a_j) \\ &= \sum_{j=1}^N p(|a_i\rangle; a_j) p(a_j) = p(a_i), \end{aligned} \quad (11)$$

where we have used Equation (10) in the conditional probabilities of Equation (11).

We note from Equations (10) and (11) that $p(|a_i\rangle) = p(|a_i\rangle, a_i)$ and that $\{|a_i\rangle\}$ represent a set of complete and mutually exclusive events:

$$p(|a_i\rangle, |a_j\rangle) = \delta_{ij}, \quad \sum_{i=1}^N p(|a_i\rangle) = 1. \quad (12)$$

Lemma 1 can be easily generalized for entangled states as well. For instance, consider the state:

$$|\psi_{SB}\rangle = \sum_{i=1}^N \sum_{k=1}^M c_{ik} |a_i\rangle |b_k\rangle, \quad (13)$$

where \hat{B} is a new observable commuting with \hat{A} , with eigenvalues $\{b_k\}$ and eigenstates $\{|b_k\rangle\}$, $k = 1, \dots, M$, and $|\psi_{SB}\rangle \in \mathcal{H}^N \otimes \mathcal{H}^M$, (we point out that \hat{B} does not necessarily have to belong to another system: for example, \hat{A} could be the spin observable and \hat{B} the position observable of one particle). Consider the pairs of eigenvalues $\{a_i, b_k\}$, and eigenstates $\{|a_i\rangle, |b_k\rangle\}$. By using conditional probabilities, it follows that on the generic state $|\psi_{SB}\rangle$:

$$\begin{aligned} p(|a_i\rangle | b_k\rangle) &= p(|a_i\rangle, |b_k\rangle) \\ &= p(\{a_i, b_k\}) \end{aligned} \quad (14)$$

(see Appendix A for a proof of this). Then the probability $p(a_i)$ is again given by:

$$\begin{aligned} p(a_i) &= \sum_{k=1}^M p(\{a_i, b_k\}) = \sum_{k=1}^M p(|a_i\rangle | b_k\rangle) \\ &= \sum_{k=1}^M p(\{|a_i\rangle, |b_k\rangle\}) = p(|a_i\rangle). \end{aligned} \quad (15)$$

With a similar reasoning we can also show that $p(b_k) = p(|b_k\rangle)$. We notice from Equations (14) and (15) that the outcomes of $\{|a_i\rangle, |b_k\rangle\}$ are mutually exclusive events as well, with the condition of normalization following from the normalization of $p(a_i)$.

Lemma 2: If a coefficient of an eigenstate in Equation (7) is zero, the probability of obtaining the corresponding eigenvalue and eigenstate is zero as well.

Proof: If, for instance, $c_k = 0$, the corresponding state $|a_k\rangle$ will be orthogonal to $|\psi_S\rangle$. We can think $\{|\psi_S\rangle, |a_k\rangle\}$ as eigenstates of the projector $\hat{\Pi}_S = |\psi_S\rangle\langle\psi_S|$ with eigenvalues $\{1, 0\}$. The Hermitian operator $\hat{\Pi}_S$ represents the observable testing whether the system \mathcal{S} is prepared with the state $|\psi_S\rangle$ or not. The pair $\{|\psi_S\rangle, |a_k\rangle\}$ is a complete orthonormal basis of a subspace \mathcal{H}^2 of \mathcal{H}^N .

With the system prepared in the state $|\psi_S\rangle$: $p(|\psi_S\rangle; |\psi_S\rangle) = 1$. From *Lemma 1*, it follows that $p(1; |\psi_S\rangle) = 1$ and because eigenvalues are exclusive

events $p(0; |\psi_S\rangle) = 0$. So, if we apply *Lemma 1* again to the previous equation, considering that $|a_k\rangle$ is eigenstate of both $\hat{\Pi}_S$ and \hat{A} , we have $p(|a_k\rangle; |\psi_S\rangle) = 0$ and $p(a_k; |\psi_S\rangle) = 0$.

Also *Lemma 2* can be directly generalized to entangled states. For instance, if in the state $|\psi_{SB}\rangle$ the coefficient $c_{ki} = 0$ then $p(|a_k\rangle | b_i\rangle; |\psi_{SB}\rangle) = 0$. In particular, if the state of the system is in a Schmidt decomposition:

$$|\psi_{SB}\rangle = \sum_{i=1}^N c_i |a_i\rangle |b_i\rangle, \quad (16)$$

from *Lemma 1* and *Lemma 2*, since $p(|a_i\rangle | b_k\rangle) = 0$ for $k \neq i$, it follows the perfect correlation property $p(|a_i\rangle) = p(|b_i\rangle)$:

$$\begin{aligned} p(a_i) &= \sum_{k=1}^M p(\{a_i, b_k\}) = p(|a_i\rangle | b_i\rangle) = p(|a_i\rangle) \\ &= p(|b_i\rangle) = p(b_i). \end{aligned} \quad (17)$$

Lemma 3:

If the measurement is noncontextual then the probability of obtaining the eigenvalue a_i can only depend on the modulus of the coefficient c_i :

$$p(a_i) = p(|c_i|). \quad (18)$$

Proof:

Let us recall the meaning of NCP with an example (we refer to the work by Spekkens for an extensive analysis of the contextuality concept in arbitrary operational theories and in QM [17]). Consider two operators \hat{A} and \hat{A}' which do not commute, $[\hat{A}, \hat{A}'] \neq 0$, in a Hilbert space with $\dim \mathcal{H} = 3$. Eigenvalues and eigenstates of \hat{A} and \hat{A}' are, respectively:

$$\begin{aligned} \hat{A}: & \{a_1, a_2, a_3\}, \{ |a_1\rangle, |a_2\rangle, |a_3\rangle \}, \\ \hat{A}': & \{a'_1, a'_2, a'_3\}, \{ |a'_1\rangle, |a'_2\rangle, |a'_3\rangle \}, \end{aligned} \quad (19)$$

where they have in common the eigenvalue a_1 and the associate eigenstate (we see that this is possible, if \hat{A} and \hat{A}' must not commute, just for $N \geq 3$). We can write the state $|\psi_S\rangle$ in the two different bases of Equation (19):

$$\begin{aligned} |\psi_S\rangle &= c_1 |a_1\rangle + c_2 |a_2\rangle + c_3 |a_3\rangle \\ &= c_1 |a_1\rangle + c'_2 |a'_2\rangle + c'_3 |a'_3\rangle. \end{aligned} \quad (20)$$

In general, the probability of obtaining the eigenvalue a_1 could depend on whether we measure the observable \hat{A} or \hat{A}' . In other words, the probability $p(a_1)$ could be different if we were in the experimental context which is prepared for measuring which value \hat{A} assumes of the triplet $\{a_1, a_2, a_3\}$ or which value \hat{A}' assumes of the triple $\{a'_1, a'_2, a'_3\}$.

With the result of *Lemma 1* and the completeness

postulate of QM, we can write the more general probability in different experimental contexts as:

$$\begin{aligned}
 p_A(a_1) &= p(|a_1\rangle; |\psi_s\rangle, \{|a_1\rangle, |a_2\rangle, |a_3\rangle\}), \\
 p_{A'}(a_1) &= p(|a_1\rangle; |\psi_s\rangle, \{|a_1\rangle, |a'_2\rangle, |a'_3\rangle\}).
 \end{aligned}
 \tag{21}$$

The probability is noncontextual, and we have measurement noncontextuality, if $p(a_1)$ does not depend on the measurement with other eigenvalues:

$$p_A(a_1) \equiv p_{A'}(a_1), \tag{22}$$

or, according to Lemma 1, the probability $p(|a_1\rangle)$ does not depend on the measurement with other eigenstates:

$$p_A(|a_1\rangle) \equiv p_{A'}(|a_1\rangle). \tag{23}$$

Then, by assuming NCP, Equations (21) can be reduced to the equation:

$$p(a_1) = p(|a_1\rangle; |\psi_s\rangle), \tag{24}$$

stating that the probability only depends on the eigenstate $|a_1\rangle$ and, from the completeness postulate of QM and NCP, on the state of the system $|\psi_s\rangle$.

Thus, let us choose a new basis $\{|a'_i\rangle\}$ such that the state in Equation (7) is written with only two components:

$$|\psi_s\rangle = c_1 |a_1\rangle + |c'_2\rangle |a'_2\rangle, \tag{25}$$

where $c'_i = 0$ for $i \geq 3$, $|c'_2|$ is the modulus of the amplitude c'_2 and

$$|a'_1\rangle \equiv |a_1\rangle, \quad |a'_2\rangle = \frac{1}{|c'_2|} \sum_{i=2}^N c_i |a_i\rangle. \tag{26}$$

$\{|a_1\rangle, |a'_2\rangle\}$ could be regarded as eigenstates of an observable $\hat{A}' = |a_1\rangle\langle a_1|$ with eigenvalues $\{1, 0\}$ checking whether our system is in the state $|a_1\rangle$ or not.

According to NCP the transformation of basis $\{|a_i\rangle\} \rightarrow \{|a'_i\rangle\}$ of the state cannot change the probability:

$$\begin{aligned}
 p(|a_1\rangle; |\psi_s\rangle) &= p(|a_1\rangle; |\psi_s\rangle)_{\{|a_i\rangle\}} \\
 &= p(|a_1\rangle; |\psi_s\rangle)_{\{|a'_i\rangle\}}.
 \end{aligned}
 \tag{27}$$

Now we expressly write the phase of the coefficient c_1 in the state of Equation (25):

$$|\psi_s\rangle = e^{i\phi_1} |c_1| |a_1\rangle + |c'_2| |a'_2\rangle. \tag{28}$$

We get for the state $|a_1\rangle$ the probability:

$$p(|a_1\rangle; |\psi_s\rangle) = p(|a_1\rangle; e^{i\phi_1} |c_1| |a_1\rangle + |c'_2| |a'_2\rangle). \tag{29}$$

Therefore, considering the normalization $|c_1|^2 + |c'_2|^2 = 1$ and that the components of the orthonormal basis are the constants $\langle a'_i | a'_j \rangle = \delta_{ij}$, the most general function of the probability in Equation (29) can be written as:

$$p(|a_1\rangle; |\psi_s\rangle) = p(\phi_1, |c_1|). \tag{30}$$

A similar argument when applied to every state $|a_i\rangle$ gives

$$p(|a_i\rangle; |\psi_s\rangle) = p(\phi_i, |c_i|). \tag{31}$$

However, by using the conservation of probability we can show the independence from the phase of $p(|a_i\rangle; |\psi_s\rangle)$. Indeed, let us suppose to have two different states $|\psi_s\rangle$ and $|\psi_s^*\rangle$ which are distinct just with respect to the phase of the component $|a_1\rangle$ (to simplify the notation, we select a basis such that the components of the states are different from zero only in a subspace with $\dim \mathcal{H} = 2$):

$$\begin{aligned}
 |\psi_s\rangle &= e^{i\phi_1} |c_1| |a_1\rangle + e^{i\phi_2} |c_2| |a_2\rangle, \\
 |\psi_s^*\rangle &= e^{i\phi_1^*} |c_1| |a_1\rangle + e^{i\phi_2} |c_2| |a_2\rangle,
 \end{aligned}
 \tag{32}$$

where with ϕ_1^* we have denoted the new phase. The conservation of the probability requires:

$$\begin{aligned}
 p(|a_1\rangle; |\psi_s\rangle) + p(|a_2\rangle; |\psi_s\rangle) &= 1, \\
 p(|a_1\rangle; |\psi_s^*\rangle) + p(|a_2\rangle; |\psi_s^*\rangle) &= 1,
 \end{aligned}
 \tag{33}$$

where from Lemma 2 we have used the fact that $p(|a_i\rangle; |\psi_s\rangle) = 0$ if $|a_i\rangle \perp |\psi_s\rangle$. Thanks to Equation (31), we have that

$$p(|a_2\rangle; |\psi_s\rangle) = p(|a_2\rangle; |\psi_s^*\rangle). \tag{34}$$

From this and Equation (33), we also get:

$$p(|a_1\rangle; |\psi_s\rangle) = p(|a_1\rangle; |\psi_s^*\rangle), \tag{35}$$

which is equivalent to write:

$$p(\phi_1^*, |c_1|) = p(\phi_1, |c_1|). \tag{36}$$

Therefore, since ϕ_1^* is an arbitrary phase, the probability $p(a_1)$ does not depend on it but can only depend on the modulus $|c_1|$. So, we can set $p(a_1) = p(|c_1|)$, and this proves our lemma (the generalization to probabilities of other eigenvalues a_i is immediate).

Now we have all the ingredients to prove our

Theorem 1:

the probability $p(a_i)$ is given by the square modulus of coefficient c_i .

Proof:

Let us consider the state $|\psi_s\rangle \in \mathcal{H}^N$ with $N \geq 3$ in Equations (7). For the sake of simplicity we focus our attention again on $p(a_1)$. We choose two bases $\{|a_i\rangle\}$ and $\{|a'_i\rangle\}$ such that

$$|\psi_s\rangle = c_1 |a_1\rangle + c_2 |a_2\rangle + c_3 |a_3\rangle \tag{37}$$

$$= c_1 |a_1\rangle + c'_2 |a'_2\rangle, \tag{38}$$

where $c_i = 0$ for $i \geq 4$, $c'_i = 0$ for $i \geq 3$ and

$|a_1\rangle = |a'_1\rangle$. From the conservation of the total probability for the states of the two different bases we get respectively:

$$p(|a_1\rangle) + p(|a_2\rangle) + p(|a_3\rangle) = 1, \tag{39}$$

$$p(|a_1\rangle) + p(|a'_2\rangle) = 1, \tag{40}$$

where from NCP we have the same probability $p(|a_1\rangle)$ in both Equations (39) and (40). With a change of variable we can write Equation (18) as $p(a_i) = f(|c_i|^2)$. From Equations (39) and (40) it follows that

$$\begin{aligned} f(|c_1|^2) + f(|c_2|^2) + f(|c_3|^2) &= 1, \\ f(|c_1|^2) + f(|c'_2|^2) &= 1. \end{aligned} \tag{41}$$

By comparing the two Equations (41) we get:

$$f(|c'_2|^2) = f(|c_2|^2) + f(|c_3|^2). \tag{42}$$

Since the state $|\psi_s\rangle$ is normalized, we also have

$$|c_1|^2 + |c_2|^2 + |c_3|^2 = 1, \tag{43}$$

$$|c_1|^2 + |c'_2|^2 = 1, \tag{44}$$

and

$$|c'_2|^2 = |c_2|^2 + |c_3|^2. \tag{45}$$

Putting this in Equation (42):

$$f(|c_2|^2 + |c_3|^2) = f(|c_2|^2) + f(|c_3|^2). \tag{46}$$

Now we recall that a function $f(x)$ is linear with respect to the variable x if and only if the two following properties are jointly satisfied:

$$\begin{aligned} f(x_1 + x_2) &= f(x_1) + f(x_2) \quad \text{additivity,} \\ f(\alpha x) &= \alpha f(x) \quad \text{homogeneity,} \end{aligned} \tag{47}$$

where α is any real number. We remark that if α is a rational number, it can be shown that the homogeneity follows from the additivity. Furthermore, since rational numbers form a subset dense in the set of real numbers, such derivation can be extended to the case of irrational α , if f is a continuous function [35].

Therefore, the condition of additivity is sufficient to establish the linearity if f is a continuous function. On the contrary, if f is not assumed to be continuous, the condition of additivity implies linearity only for rational values of the variable x .

Hence, if we assume that the probability is continuous with respect to the coefficients $|c_i|^2$, Equation (46) is a constraint of linearity for the probability. In this case we can write:

$$p(a_i) = k|c_i|^2, \quad i = 2, 3, \tag{48}$$

where k denotes a constant. From Lemma 1 we have that $p(|c_i|^2 = 1) = 1$ and therefore $k = 1$ and $p(a_i) = |c_i|^2$. Hence, the Born rule is deduced for continuous probabilities.

We point out that we have postulated the continuity of probability. However, a derivation of such a property from noncontextuality is present in Gleason's work and it occupies a prominent part of that paper. Therefore in comparison to Gleason we have only accomplished half the job: as Bertrand Russell once said, postulating is equivalent to theft on the honest fatigue. On the other hand, if the continuity of probability is neither postulated nor demonstrated, we can only say that the Born rule applies only for rational values of the coefficients $|c_i|^2$.

Thus we now intend to complete our derivation by showing from NCP that the probability is continuous and that the Born rule also holds for irrational coefficients.

Consider, once again, $p(|a_1\rangle)$. We select a basis $\{|a'_i\rangle\}$ such that the state $|\psi_s\rangle$ is written with two components as in Equation (38), with $|a'_1\rangle \equiv |a_1\rangle$, and the conservation of probability is given by Equation (40).

Then we choose a new basis $\{|a''_i\rangle\}$ in general with $|a''_1\rangle \neq |a_1\rangle$, $c''_i = 0$ for $i \geq 4$, but with $|a''_2\rangle \equiv |a'_2\rangle$ so that we can write

$$|\psi_s\rangle = c''_1|a''_1\rangle + c''_2|a''_2\rangle + c''_3|a''_3\rangle. \tag{49}$$

The conservation of probability for the new states of the basis requires that

$$p(|a''_1\rangle) + p(|a''_2\rangle) + p(|a''_3\rangle) = 1. \tag{50}$$

From the previous condition we have the inequality:

$$p(|a''_1\rangle) + p(|a'_2\rangle) \leq 1. \tag{51}$$

By comparing Equation (40) with Equation (51) and since, thanks to NCP, the probability $p(|a'_2\rangle)$ is the same in these equations, we get

$$p(|a''_1\rangle) \leq p(|a_1\rangle). \tag{52}$$

Finally, we choose a third basis $\{|a'''_i\rangle\}$ such that $c'''_i = 0$ for $i \geq 4$, this time with $|a'''_1\rangle \equiv |a_1\rangle$:

$$|\psi_s\rangle = c'''_1|a_1\rangle + c'''_2|a'''_2\rangle + c'''_3|a'''_3\rangle. \tag{53}$$

The conservation of probability gives us:

$$p(|a_1\rangle) + p(|a'''_2\rangle) + p(|a'''_3\rangle) = 1. \tag{54}$$

From Equation (54) we get the inequality:

$$p(|a_1\rangle) \leq 1 - p(|a'''_2\rangle). \tag{55}$$

Therefore the probability $p(|a_1\rangle)$ is in the range:

$$p(|a''_1\rangle) \leq p(|a_1\rangle) \leq 1 - p(|a'''_2\rangle). \tag{56}$$

Suppose $|c_1|^2$ to be an irrational number, we choose

rational values for $|c_1^n|^2$ and $|c_2^m|^2$ so that $p(|a_1^n\rangle) = |c_1^n|^2$ and $p(|a_2^m\rangle) = |c_2^m|^2$. By putting $p(|a_1\rangle) = p(|c_1|^2)$, Equation (56) becomes

$$|c_1^n|^2 \leq p(|c_1|^2) \leq 1 - |c_2^m|^2, \quad |c_1^n|^2, |c_2^m|^2 \in \mathbb{Q}, \quad (57)$$

with $|c_1^n|^2 \leq 1 - |c_2^m|^2$. Let us denote with $|\psi_{s\perp}\rangle$ the orthonormal state to $|\psi_s\rangle$. Consider the limits $|a_1^n\rangle \rightarrow |a_1\rangle$, $|a_3^m\rangle \rightarrow |\psi_{s\perp}\rangle$ and $|a_2^m\rangle \rightarrow |a_2'\rangle$, $|a_3^m\rangle \rightarrow |\psi_{s\perp}\rangle$. From

Equation (38) we have $|c_1^n|^2 \rightarrow |c_1|^2$, $|c_3^m|^2 \rightarrow 0$ and $|c_2^m|^2 \rightarrow |c_2'|^2$, $|c_3^m|^2 \rightarrow 0$. Summarizing, with the normalization condition in Equation (44) we get:

$$|c_1^n|^2 \rightarrow |c_1|^2, \quad 1 - |c_2^m|^2 \rightarrow |c_1|^2. \quad (58)$$

Because rational numbers form a dense set in the set of real numbers, we can always find a couple of numbers $|c_1^n|^2, 1 - |c_2^m|^2 \in \mathbb{Q}$, which are as near as we want to the irrational number $|c_1|^2$. Therefore, from Equation (57) it follows that the oscillation of the function p is zero at $|c_1|^2$. Consequently, the probability must be continuous at $|c_1|^2$ and $p(|c_1|^2) = |c_1|^2$ even if $|c_1|^2 \notin \mathbb{Q}$.

3. Noncontextual Probability for Entangled States

Theorem 1 with NCP can be generalized to include also entangled states, bipartite or multipartite systems and quantum measurement devices.

Let us indicate with \mathcal{C} the context, including our experimental devices and the environment, and with $|C_i\rangle \in \mathcal{H}^M$ its possible states. Consider again the measurement of the two observables \hat{A} and \hat{A}' introduced in the proof of *Lemma 3* in the previous section. The choice of measuring either the eigenvalues $\{a_2, a_3\}$ or $\{a_2', a_3'\}$ together with a_1 corresponds to two different experimental arrangements \mathcal{A} and \mathcal{A}' .

Suppose that at time t_0 the system is in the initial state $|\psi_{sc}(t_0)\rangle = |\psi_s\rangle|C_0\rangle$. At time t , after the interaction with the measurement device, the entanglement between the system \mathcal{S} with the experimental setup \mathcal{C} correlates the observable \hat{A} (or \hat{A}') of the system with the pointers of the experimental device. If we measure \hat{A} we have the state:

$$|\psi_{sc}(t)\rangle = c_1|a_1\rangle|C_1\rangle + c_2|a_2\rangle|C_2\rangle + c_3|a_3\rangle|C_3\rangle, \quad (59)$$

while if we measure \hat{A}' we get:

$$|\psi_{sc'}(t)\rangle = c_1|a_1\rangle|C_1'\rangle + c_2'|a_2'\rangle|C_2'\rangle + c_3'|a_3'\rangle|C_3'\rangle. \quad (60)$$

The global wave functions Equations (59) and (60) are different in the different contexts \mathcal{A} and \mathcal{A}' . So the two probabilities

$$\begin{aligned} p_A(a_1) &= p(|a_1\rangle|C_1\rangle; |\psi_{sc}(t)\rangle), \\ p_{A'}(a_1) &= p(|a_1\rangle|C_1'\rangle; |\psi_{sc'}(t)\rangle), \end{aligned} \quad (61)$$

may be different. The possible contextuality of probability is now included inside the description of the global system $\mathcal{S} + \mathcal{C}$ of the quantum state. We notice that it is no longer possible to write

$$p_A(a_1) = p(|a_1\rangle|C_1\rangle; |\psi_{sc}(t)\rangle, \{|a_1\rangle, |a_2\rangle, |a_3\rangle\}), \quad (62)$$

if we accept the completeness of QM. Otherwise, indeed, some parameters would be outside the description of the state $|\psi_{sc}(t)\rangle$, which now englobes not only the system but also the context of the measurement. This should imply the existence of hidden variables.

Assuming NCP requires the equality between the probabilities in Equation (61):

$$p(|a_1\rangle|C_1\rangle; |\psi_{sc}(t)\rangle) = p(|a_1\rangle|C_1'\rangle; |\psi_{sc'}(t)\rangle), \quad (63)$$

which generalizes the condition in Equation (23). In fact, Equation (63), by the perfect correlation property of Equation (17), becomes equivalent to Equation (23).

In a similar way as in the previous derivation, by using the state of Equation (59) and a state $|\psi_s\rangle$ with a new basis $\{|a_i'\rangle\}$ such that $c_i' = 0$ for $i \geq 3$, $|a_1'\rangle \equiv |a_1\rangle$ we can write:

$$|\psi_{sc'}(t)\rangle = c_1|a_1\rangle|C_1'\rangle + c_2'|a_2'\rangle|C_2'\rangle. \quad (64)$$

From the condition of NCP it is not difficult to deduce that $p(|a_i\rangle|C_i\rangle; |\psi_{sc}(t)\rangle) = p(|c_i|)$ must hold. Then, employing states like in Equation (59) and Equation (64) and from NCP, normalization condition and conservation of probability, it is possible to prove again the linearity of the probabilities with respect to the square moduli of the amplitudes (continuity can also be generalized to the present case). Putting all pieces together, we get

$$\begin{aligned} p(|a_i\rangle|C_i\rangle; |\psi_{sc}(t)\rangle) &= |c_i|^2 \\ &= |\langle a_i | \langle C_i | |\psi_{sc}(t)\rangle|^2 \\ &= |\langle a_i | \psi_s \rangle|^2. \end{aligned} \quad (65)$$

We note that the state of the global system $\mathcal{S} + \mathcal{C}$ is in the Hilbert space $\mathcal{H}^N \otimes \mathcal{H}^M$ with dimension $N \cdot M$. If $M \geq 3$ this allows us to include in the derivation of the Born rule also states with $N = 2$. In fact, consider a system \mathcal{S} described by a state $|\psi_s\rangle$ in a Hilbert space with $\dim \mathcal{H} = 2$ and basis $\{|a_1\rangle, |a_2\rangle\}$. With a measurement we could use the context \mathcal{C} as an ancilla system to build the following entangled states:

$$\begin{aligned} |\psi_{sc}^0\rangle &= c_1|a_1\rangle|C_1\rangle + c_2|a_2\rangle|C_2\rangle, \\ |\psi_{sc'}^1\rangle &= c_1|a_1\rangle|C_1'\rangle + c_2'|a_2\rangle|C_2'\rangle + c_3'|a_2\rangle|C_3'\rangle, \end{aligned} \quad (66)$$

where $|\psi_{SC'}^1\rangle$ is obtained by dividing the part of the state $|\psi_S\rangle$ associated with the eigenvalue a_2 into two regions. So it is possible to demonstrate that $p(a_i) = |\langle a_i | \psi_S \rangle|^2$ as in the previous cases. We note that the condition $\dim \mathcal{H} \geq 3$ for the state of the ancilla system \mathcal{C} coupled with \mathcal{S} is necessary to obtain, from Equations (66) by the NCP property of Equation (63), a constraint of additivity and linearity similar to that which is at Equations (39)-(46).

4. Noncontextual Probability and Resolution of the Measurement

A simple application of Lemma 3 shows that equal amplitudes are related to equal probabilities, i.e: if $c_i = c_k$ with $i \neq k$ then $p(a_i) = p(a_k)$. Whereas Lemma 3 is derived from the NCP principle, the same result was obtained by Zurek and Deutsch on the base of two different invariance principles. Then both authors study the more general case with states having different rational amplitudes using the fine-graining technique. We shortly summarize their results. We consider the fine-graining in Zurek's approach with the environment in a simple situation (our remarks also hold for Deutsch's similar derivation). We assume that the state of the system \mathcal{S} is in a Hilbert space with $\dim \mathcal{H} = 2$, entangled with the state of the environment \mathcal{E} such that their joint state is

$$|\psi_{SE}\rangle = \frac{1}{\sqrt{N}}|a_1\rangle|\varepsilon_1\rangle + \sqrt{\frac{N-1}{N}}|a_2\rangle|\varepsilon_2\rangle. \quad (67)$$

The idea is to transform this state with unequal coefficients into a state with equal coefficients by fine-graining. This goal is reached by extracting from the environment an ancilla system \mathcal{C} having states

$$|C_1\rangle = |C'_1\rangle, \quad |C_2\rangle = \sum_{i=2}^N |C'_i\rangle / \sqrt{N-1}, \quad (68)$$

which become correlated with the states $\{|a_1\rangle, |a_2\rangle\}$ of \mathcal{S} . Within a quantum measurement, \mathcal{C} can be regarded as a counter and $\{|C'_i\rangle\}$ as a new orthonormal basis of pointer states. Denoting with $\{|e_i\rangle\}$ the new states of the environment without \mathcal{C} , after the interaction between \mathcal{S} and \mathcal{C} we have the joint system

$$|\psi_{SC\mathcal{E}}\rangle = \frac{1}{\sqrt{N}}|a_1\rangle|C_1\rangle|e_1\rangle + \sqrt{\frac{N-1}{N}}|a_2\rangle|C_2\rangle|e_2\rangle, \quad (69)$$

which can be expressed with the new basis of pointer states of Equation (68):

$$|\psi_{SC\mathcal{E}}\rangle = \frac{1}{\sqrt{N}}|a_1\rangle|C_1\rangle|e_1\rangle + \frac{1}{\sqrt{N}}\sum_{i=2}^N |a_2\rangle|C'_i\rangle|e_2\rangle. \quad (70)$$

Now we have a new state with equal amplitudes. But how can we be sure that after fine-graining the probability $p(|a_1\rangle)$ with the state $|\psi_{SC\mathcal{E}}\rangle$ is the same

with the state $|\psi_{SC\mathcal{E}}\rangle$ or $|\psi_{SE}\rangle$? In principle they could be different, so to legitimately use the fine-graining reasoning we have to assume the equality:

$$p(|a_1\rangle; |\psi_{SC\mathcal{E}}\rangle) = p(|a_1\rangle; |\psi_{SE}\rangle), \quad (71)$$

equivalent to the condition of NCP encountered in Equation (63) of Section 3. Then, in their derivations Zurek and Deutsch implicitly assume NCP. We notice that fine-graining is equivalent to a change of measurement resolution, and this has been used in our derivation as well. A change of resolution, for instance, is considered going from Equation (39) to Equation (40). There we assumed from NCP that $p(|a_1\rangle)$ is the same in the two equations and therefore:

$$p(|a'_2\rangle) = p(|a_2\rangle) + p(|a_3\rangle), \quad (72)$$

where $p(|a'_2\rangle)$ is the probability of finding the state of the system in the subspace orthonormal to $|a_1\rangle$, whereas $p(|a_2\rangle)$ and $p(|a_3\rangle)$ are the probabilities of finding it, respectively, in the subspaces $|a_2\rangle$ and $|a_3\rangle$ of the subspace orthonormal to $|a_1\rangle$. In fact, in terms of projectors, Equation (72) can be written as Gleason's Postulate 2:

$$\begin{aligned} & p(|a_2\rangle\langle a_2| + |a_3\rangle\langle a_3|) \\ &= p(|a_2\rangle\langle a_2|) + p(|a_3\rangle\langle a_3|). \end{aligned} \quad (73)$$

Thus NCP could be physically interpreted as a probability invariance when we change the resolution of the measurement connected with other properties of the system. This interpretation gives physical insight to the NCP principle as an invariance under changing resolution. The probability of one measurement outcome should not change if we make coarse or fine-graining over other measurement outcomes. We point out that coarse/fine graining could be done through a post-processing after the first measurement. Therefore, NCP is a natural assumption if we do not want signalling from the future to the past.

5. Noncontextual Probability and No-Faster-than-Light Signalling

In general, every measurement in the real world happens in the spacetime. Every measurement of any observable is necessarily reduced to a measurement of position at a certain time. For instance, the Stern-Gerlach apparatus correlates a spin observable with a position observable so that different values of the spin univocally correspond to different position values of the system. A measurement of position detects the spin value of the system.

Let us introduce in our representation of \mathcal{S} the observable position \hat{X} . We assume that at time t_0 , the system is in the initial state $|\psi_{SC}(t_0)\rangle = |\psi_S\rangle|x_0\rangle|C_0\rangle$.

For simplicity of notation, from now on we denote with $|C_{xi}\rangle = |x_i\rangle|C_i\rangle$ the position state of the system with its context, where the letter x will remind us the dependence on the position of the system and the experimental apparatus. Suppose that we can measure either the values of the observable \hat{A} or the values of the observable \hat{A}' . In the context A of measurement we get:

$$|\psi_{SC}(t)\rangle = c_1|a_1\rangle|C_{x1}\rangle + c_2|a_2\rangle|C_{x2}\rangle + c_3|a_3\rangle|C_{x3}\rangle, \quad (74)$$

while for the experimental context A' we have the different state:

$$|\psi_{SC'}(t)\rangle = c'_1|a'_1\rangle|C'_{x1}\rangle + c'_2|a'_2\rangle|C'_{x2}\rangle + c'_3|a'_3\rangle|C'_{x3}\rangle. \quad (75)$$

We want to prove the following

Theorem 2:

If we assume the no-faster-than-light signalling condition for measurement events which are spacelike separate, then the probability $p(a_i)$ must be non-contextual.

Proof:

Consider a set of identically prepared systems \mathcal{S} in the initial state $|\psi_{SC}(t_0)\rangle$. In order to estimate the probabilities $\{p(a_1), p(a_2), p(a_3)\}$ and $\{p(a'_1), p(a'_2), p(a'_3)\}$, let us assume that our experimental device is arranged to measure how many systems have the eigenvalue a_1 at the point x_1 and the eigenvalues $\{a_2, a_3\}$ (Equation (74)) or the eigenvalues $\{a'_2, a'_3\}$ (Equation (75)) at the points x_2 and x_3 . This could be done by first dividing the systems \mathcal{S} in two parts, one associated with the state $|a_1\rangle$ at the point x_1 and an other one associated with the orthogonal state $|a_{1\perp}\rangle$ at a different point $x_{2\perp}$ (see the superposition state in Equation (78)), then by dividing such a state in two sub-parts, associated either with the basis $\{|a_2\rangle, |a_3\rangle\}$, or with the basis $\{|a'_2\rangle, |a'_3\rangle\}$ at the points x_2, x_3 , respectively connected with the states in Equations (74)-(75).

If the measurements of $\{a_2, a_3\}$ or $\{a'_2, a'_3\}$ at the points x_2 and x_3 are events spacelike separate with respect to the measurement of a_1 at x_1 , according to the relativistic hypothesis of no-signalling, the probability $p(a_1)$ of getting the eigenvalue a_1 has to be the same in both different experimental contexts A and A' . Otherwise an observer situated in the place of measurement of $\{a_2, a_3\}$ or $\{a'_2, a'_3\}$ would be able to send a superluminal signal to another observer situated in the place where a_1 is measured. Then, according to the no-faster-than-light signalling hypothesis:

$$p(|a_1\rangle|C_{x1}\rangle; |\psi_{SC}(t)\rangle) = p(|a_1\rangle|C'_{x1}\rangle; |\psi_{SC'}(t)\rangle), \quad (76)$$

which is to say, with the result of Equation (17):

$$p(a_1) = p(|a_1\rangle; |\psi_{SC}(t)\rangle) = p(|a_1\rangle; |\psi_{SC'}(t)\rangle). \quad (77)$$

Therefore $p(a_1)$ must be noncontextual.

Consider, beside the measurement of the eigenvalues $\{a_1, a_2, a_3\}$ of the observable \hat{A} , also the simple possibility of measuring whether our systems assume the eigenvalue a_1 or not, which is equivalent to check whether the systems \mathcal{S} are in the state $|a_1\rangle$ or in the subspace $|a_{1\perp}\rangle$ orthogonal to it. In such a case, during the process of measurement, the systems $\mathcal{S} + \mathcal{C}$ will be in the state:

$$|\psi_{Sx_{1\perp}}(t)\rangle = c_1|a_1\rangle|C_{x1}\rangle + c_{2\perp}|a_{1\perp}\rangle|C_{x2\perp}\rangle. \quad (78)$$

If the two choices of measuring either $\{|a_1\rangle, |a_2\rangle, |a_3\rangle\}$ or $\{|a_1\rangle, |a_{1\perp}\rangle\}$ correspond to events spacelike separate, as mentioned above, from the condition of no-signalling we must have the NCP condition:

$$p(a_1) = p(|a_1\rangle; |\psi_{SC}(t)\rangle) = p(|a_1\rangle; |\psi_{Sx_{1\perp}}(t)\rangle). \quad (79)$$

From Equation (79) the conservation of probability imposes also the equation:

$$\begin{aligned} & p(|a_{1\perp}\rangle; |\psi_{Sx_{1\perp}}(t)\rangle) \\ &= p(|a_2\rangle; |\psi_{SC}(t)\rangle) + p(|a_3\rangle; |\psi_{SC}(t)\rangle). \end{aligned} \quad (80)$$

As we showed at the end of Section IV, Equation (80) is equivalent to the initial hypothesis adopted by Gleason. Moreover, by examining Equation (79) we understand that it is analogous to assuming an invariance condition of the probability under a resolution change of our measurement, or fine/coarse graining. So, at least for spacelike separate measurements, NCP could be deduced by a relativistic principle. However, the Born rule holds for general events, not only for spacelike separate ones. Hence NCP or, equivalently, the invariance of probability when we change resolution, seem to be more general principles.

6. Concluding Remarks

We have given a new derivation of the Born rule based on the noncontextuality for projective measurements and on some non-statistical postulates of QM. The introduction of a pure quantum state of the system has allowed a more elementary proof than Gleason's theorem. This may seem, at first sight, a loss of generality with respect to Gleason's proof, however Gleason also introduces vectors of basis which can be regarded as states of the system. Moreover, the step from pure states to density operators and to the trace rule for mixed ensembles is quite natural when we already have the Born rule [36].

As in Gleason's, our derivation holds for $\dim\mathcal{H} \geq 3$, where $\dim\mathcal{H}$ could be finite or infinite. The origin of this dimensional limitation lies in the fact that for $\dim\mathcal{H} \geq 3$ the same state vector may belong to several

distinct orthonormal bases. This, with NCP, the normalization of the probability and the quantum state, brings us to a constraint of functional linearity between the probability and the amplitude square modulus. Our theorem can be simply generalized in order to include multipartite states. In such a case, if a part of the global system has $\dim\mathcal{H} \geq 3$, the Born rule can also be proved to hold in subspaces with $\dim\mathcal{H} = 2$.

We have also given a derivation that the probability is continuous with respect to the amplitudes using measurement noncontextuality. This fills a gap of Zurek's and Deutsch's demonstrations.

Physical motivation for NCP could come from a relativistic principle of no-faster-than-light signalling. However, the Born rule also holds for measurements which are non-spacelike separate, hence NCP seems to be a more fundamental principle. Finally, we have given an interpretation of measurement noncontextuality as an invariance of the probability under a resolution change of our measurement.

7. Acknowledgements

This work has been supported by the EU-STREP Project QIBEC within the activities of the BEC center. QSTAR is the MPQ, IIT, LENS, UniFi joint center for Quantum Science and Technology in Arcetri.

REFERENCES

- [1] J. A. Wheeler and W. H. Zurek, Eds., "Quantum Theory and Measurement," Princeton University Press, Princeton, 1963.
- [2] J. von Neumann, "Mathematical Foundations of Quantum Mechanics," Princeton University Press, Princeton, 1955. [doi:10.1063/1.3061789](https://doi.org/10.1063/1.3061789)
- [3] T. F. Jordan, "Linear Operators for Quantum Mechanics," John Wiley Sons, New York, 1969. [doi:10.1119/1.1975255](https://doi.org/10.1119/1.1975255)
- [4] A. M. Gleason, "Measures on the Closed Subspaces of a Hilbert Space," *Journal of Mathematics and Mechanics*, Vol. 6, No. 6, 1957, pp. 885-893.
- [5] H. Everett, "Relative State Formulation of Quantum Mechanics," *Reviews of Modern Physics*, Vol. 29, No. 3, 1957, pp. 454-462. [doi:10.1103/RevModPhys.29.454](https://doi.org/10.1103/RevModPhys.29.454)
- [6] B. S. DeWitt, "The Many-Universes Interpretation of Quantum Mechanics," In: B. d'Espagnat, Ed., *Foundations of Quantum Mechanics*, Academic Press, New York, 1971.
- [7] D. Finkelstein, "The Logic of Quantum Mechanics," *Transactions of the New York Academy of Sciences*, Vol. 25, No. 6, 1965, pp. 621-637. [doi:10.1111/j.2164-0947.1963.tb01483.x](https://doi.org/10.1111/j.2164-0947.1963.tb01483.x)
- [8] J. B. Hartle, "Quantum Mechanics of Individual Systems," *American Journal of Physics*, Vol. 36, No. 8, 1968, pp. 704-712. [doi:10.1119/1.1975096](https://doi.org/10.1119/1.1975096)
- [9] N. Graham, "The Measurement of Relative Frequency," In: B. S. DeWitt and N. Graham, Eds., *The Many-Worlds Interpretation of Quantum Mechanics*, Princeton University Press, Princeton, 1973.
- [10] R. Geroch, "The Everett Interpretation," *Noûs*, Vol. 18, No. 4, 1984, pp. 617-633. [doi:10.2307/2214880](https://doi.org/10.2307/2214880)
- [11] E. Farhi, J. Goldstone and S. Gutmann, "How Probability Arises in Quantum Mechanics," *Annals of Physics*, Vol. 192, No. 2, 1989, pp. 368-382. [doi:10.1016/0003-4916\(89\)90141-3](https://doi.org/10.1016/0003-4916(89)90141-3)
- [12] H. Stein, "The Everett Interpretation of Quantum Mechanics: Many Worlds or None?" *Noûs*, Vol. 18, No. 4, 1984, pp. 635-652. [doi:10.2307/2214881](https://doi.org/10.2307/2214881)
- [13] A. Kent, "Against Many-Worlds Interpretations," *International Journal of Modern Physics A*, Vol. 5, No. 9, 1990, pp. 1745-1762. [doi:10.1142/S0217751X90000805](https://doi.org/10.1142/S0217751X90000805)
- [14] T. Endo, "Verification of Born's Rule by a Quantum Mechanical Meter," *Physics Letters A*, Vol. 308, No. 4, 2003, pp. 256-258. [doi:10.1016/S0375-9601\(03\)00083-5](https://doi.org/10.1016/S0375-9601(03)00083-5)
- [15] C. M. Caves, C. A. Fuchs and R. Schack, "Quantum Probabilities as Bayesian Probabilities," *Physical Review A*, Vol. 65, No. 2, 2002, pp. 1-6. [doi:10.1103/PhysRevA.65.022305](https://doi.org/10.1103/PhysRevA.65.022305)
- [16] M. Redhead, "Incompleteness, Nonlocality and Realism," Clarendon Press, Oxford, 1989. [doi:10.1119/1.16032](https://doi.org/10.1119/1.16032)
- [17] R. W. Spekkens, "Contextuality for Preparations, Transformations, and Unsharp Measurements," *Physical Review A*, Vol. 71, No. 5, 2005, pp. 1-17. [doi:10.1103/PhysRevA.71.052108](https://doi.org/10.1103/PhysRevA.71.052108)
- [18] W. H. Zurek, "Environment-Assisted Invariance, Entanglement, and Probabilities in Quantum Physics," *Physical Review Letters*, Vol. 90, No. 12, 2003, pp. 1-4. [doi:10.1103/PhysRevLett.90.120404](https://doi.org/10.1103/PhysRevLett.90.120404)
- [19] M. Schlosshauer and A. Fine, "On Zurek's Derivation of the Born Rule," *Foundations of Physics*, Vol. 35, No. 197, 2005, pp. 197-213. [doi:10.1007/s10701-004-1941-6](https://doi.org/10.1007/s10701-004-1941-6)
- [20] D. Deutsch, "Quantum Theory of Probability and Decisions," *Proceedings of the Royal Society of London A*, Vol. 455, No. 1988, 1999, pp. 3129-3137. [doi:10.1098/rspa.1999.0443](https://doi.org/10.1098/rspa.1999.0443)
- [21] W. H. Zurek, "Probabilities from Entanglement, Born's Rule from Envariance," *Physical Review A*, Vol. 71, 2005, pp. 1-29. [doi:10.1103/PhysRevA.71.052105](https://doi.org/10.1103/PhysRevA.71.052105)
- [22] W. H. Zurek, "Entanglement Symmetry, Amplitudes, and Probabilities: Inverting Born's Rule," *Physical Review Letters*, Vol. 106, No. 25, 2011, pp. 1-4. [doi:10.1103/PhysRevLett.106.250402](https://doi.org/10.1103/PhysRevLett.106.250402)
- [23] C. M. Caves, "Notes on Zurek's Derivation of the Quantum Probability Rule," 2005. <http://info.phys.unm.edu/caves/reports/ZurekBornDerivation.pdf>
- [24] U. Mohrhoff, "Probabilities from Envariance?" *International Journal of Quantum Information*, Vol. 2, No. 2, 2004, pp. 221-230. [doi:10.1142/S0219749904000195](https://doi.org/10.1142/S0219749904000195)
- [25] H. Barnum, "No-Signalling-Based Version of Zurek's Derivation of Quantum Probabilities: A Note on Environment-Assisted Invariance, Entanglement, and Probabili-

- ties in Quantum Physics,” 2003.
arXiv: quant-ph/0312150
- [26] H. Barnum, C. M. Caves, J. Finkelstein, C. A. Fuchs and R. Schack, “Quantum Probability from Decision Theory?” *Proceedings of the Royal Society of London A*, Vol. 456, No. 1997, 2000, pp. 1175-1182.
[doi:10.1098/rspa.2000.0557](https://doi.org/10.1098/rspa.2000.0557)
- [27] S. Saunders, “Derivation of the Born Rule from Operational Assumptions,” *Proceedings of the Royal Society of London A*, Vol. 460, No. 2046, 2004, pp. 1-18.
[doi:10.1098/rspa.2003.1230](https://doi.org/10.1098/rspa.2003.1230)
- [28] D. Wallace, “Everettian Rationality: Defending Deutsch’s approach to probability in the Everett interpretation,” *Studies in History and Philosophy of Modern Physics*, Vol. 34, No. 3, 2003, pp. 415-438.
[doi:10.1016/S1355-2198\(03\)00036-4](https://doi.org/10.1016/S1355-2198(03)00036-4)
- [29] W. H. Zurek, “Decoherence, Einselection, and the Quantum Origins of the Classical,” *Reviews of Modern Physics*, Vol. 75, No. 3, 2003, pp. 715-775.
[doi:10.1103/RevModPhys.75.715](https://doi.org/10.1103/RevModPhys.75.715)
- [30] W. H. Zurek, “Quantum Darwinism and Envariance,” 2003.
arXiv: quant-ph/0308163
- [31] C. Cohen-Tannoudji, B. Diu and F. Laloë, “Quantum Mechanics,” Hermann and Wiley, Paris, 1977.
- [32] N. Zettili, “Quantum Mechanics, Concepts and Applications,” Wiley, Chichester, 2001.
- [33] W. Greiner, “Quantum Mechanics, An Introduction,” Springer-Verlag, Berlin, 2000.
- [34] A. Galindo and P. Pascual, “Quantum Mechanics I,” Springer-Verlag, Berlin, 1990.
[doi:10.1007/978-3-642-83854-5_2](https://doi.org/10.1007/978-3-642-83854-5_2)
- [35] M. Kuczma, “An Introduction to the Theory of Functional Equations and Inequalities,” Birkhäuser, Basel, 2009.
[doi:10.1007/978-3-7643-8749-5](https://doi.org/10.1007/978-3-7643-8749-5)
- [36] J. J. Sakurai, “Modern Quantum Mechanics,” Addison-Wesley, San Francisco, 1994.

Appendix A

We want to generalize the result of *Lemma 1* for entangled states. Suppose that the system is prepared with eigenvalues $\{a_i, b_k\}$ and therefore with eigenstates $\{|a_i\rangle, |b_k\rangle\}$. Similarly to the unentangled case, we can deduce that $p(\{a_n, b_m\}) = \delta_{in} \delta_{km}$ if and only if $p(\{|a_n\rangle, |b_m\rangle\}) = \delta_{in} \delta_{km}$. This relation, translated into conditional probabilities, is equivalent to writing:

$$\begin{aligned} p(\{a_i, b_k\}; \{|a_n\rangle, |b_m\rangle\}) \\ = p(\{|a_n\rangle, |b_m\rangle\}; \{a_i, b_k\}) = \delta_{in} \delta_{km}. \end{aligned} \quad (81)$$

Then, from Equation (81) we have that for a generic state as in Equation (13):

$$\begin{aligned} p(\{a_i, b_k\}) &= \sum_{n=1}^N \sum_{m=1}^M p(\{a_i, b_k\}, \{|a_n\rangle, |b_m\rangle\}) \\ &= \sum_{n=1}^N \sum_{m=1}^M p(\{a_i, b_k\}; \{|a_n\rangle, |b_m\rangle\}) p(\{|a_n\rangle, |b_m\rangle\}) \quad (82) \\ &= p(\{|a_i\rangle, |b_k\rangle\}). \end{aligned}$$

If the observable \hat{A} has an eigenstate $|a_i\rangle$ and the observable \hat{B} has an eigenstate $|b_k\rangle$ then the tensorial product $|a_i\rangle|b_k\rangle$ is an eigenstate of the observable $\hat{A} \otimes \hat{B}$, and vice versa. The eigenstates $\{|a_i\rangle|b_k\rangle\}$ of $\hat{A} \otimes \hat{B}$ form a basis of Hilbert space $\mathcal{H}^N \otimes \mathcal{H}^M$.

If we indicate with $p(|a_n\rangle|b_m\rangle)$ the probability of obtaining the eigenstate $|a_n\rangle|b_m\rangle$, with a measurement of $\hat{A} \otimes \hat{B}$, we have $p(\{|a_n\rangle, |b_m\rangle\}) = \delta_{in} \delta_{km}$ if and only if $p(|a_n\rangle|b_m\rangle) = \delta_{in} \delta_{km}$, or, in equivalence:

$$\begin{aligned} p(\{|a_i\rangle, |b_k\rangle\}; |a_n\rangle|b_m\rangle) \\ = p(|a_n\rangle|b_m\rangle; \{|a_i\rangle, |b_k\rangle\}) = \delta_{in} \delta_{km}. \end{aligned} \quad (83)$$

From Equations (82)-(83), through the use of the conditional probabilities, it immediately follows that on the state $|\psi_{SB}\rangle$ we have $p(|a_i\rangle|b_k\rangle) = p(\{a_i, b_k\})$.