Quantum probabilities from combination of Zurek's envariance and Gleason's theorem

Nenashev

June 1, 2014

Abstract

The quantum-mechanical rule for probabilities, in its most general form of positive-operator valued measure (POVM), is shown to be a consequence of the environment-assisted invariance (envariance) idea suggested by Zurek [Phys. Rev. Lett. 90, 120404 (2003)], being completed by Gleason's theorem. This provides also a method for derivation of the Born rule.

1. Introduction

Almost all textbooks on quantum mechanics consider only measurements of a special kind—namely, measurements of *observables*. An observable $\mathcal{O}$ corresponds to some Hermitian operator $\hat{O}$. A measuring device measures the observable $\mathcal{O}$, if (i) the only possible results of measurements are eigenvalues of $\mathcal{O}$, and (ii) if the state $|\psi\rangle$ of the system under measurement is an eigenstate of $\hat{O}$, one can predict the measurement result *with certainty* to be the eigenvalue of $\hat{O}$ corresponding to $|\psi\rangle$. When $|\psi\rangle$ is not an eigenvector of $\hat{O}$, the measurement result cannot be known in advance, but the postulates of quantum mechanics allow one to predict the *probabilities* of results. Namely, the probability $p_\lambda(|\psi\rangle)$ of the result λ is

$$p_\lambda(|\psi\rangle) = \langle\psi|\, \hat{P}_\lambda \,|\psi\rangle \tag{1}$$

where $\hat{P}_\lambda$ is the projector onto the eigenspace of $\hat{O}$ corresponding to the eigenvalue λ. In the simplest case of non-degenerate eigenvalue λ, the projector $\hat{P}_\lambda$ is equal to $|\varphi_\lambda\rangle\langle\varphi_\lambda|$, where $|\varphi_\lambda\rangle$ is the eigenvector, and Eq. (1) turns into the Born rule:

$$p_\lambda(|\psi\rangle) = \langle\psi\,|\,\varphi_\lambda\rangle\,\langle\varphi_\lambda\,|\,\psi\rangle \equiv |\,\langle\varphi_\lambda\,|\,\psi\rangle\,|^2 \tag{2}$$

Unlike this special class of measurements, *general measurements* are not associated with any observables, and probabilities of their results do not obey literally Eq. (1). Instead, the probability $p_\lambda(|\psi\rangle)$ of some result λ of a general measurement can be expressed as

$$p_\lambda(|\psi\rangle) = \langle\psi|\, \hat{A}_\lambda \,|\psi\rangle \tag{3}$$

where $\hat{A}_\lambda$ is some Hermitian operator (not necessary a projector). The set of operators $\{\hat{A}_\lambda\}$ obeys the following requirements, which are consequences of properties of probability:

(1) eigenvalues of operators $\hat{A}_\lambda$ are bound within the range [0,1];

(2) the sum $\sum_\lambda \hat{A}_\lambda$ (over all measurement results λ) is equal to the identity operator.

The set $\{\hat{A}_\lambda\}$ satisfying these requirements is usually called a *positive-operator valued measure* (POVM) [1, 2].

Such general measurements, described by POVMs via Eq. (3), occur in various contexts: as *indirect measurements*, when a system A (to be measured) first interacts with another quantum system B, and actual measurement is then performed on the system B [1-3]; as *imperfect measurements*, where a result of a measurement is subjected to a random error [3, 4]; as *continuous* and *weak measurements* [5], etc.

In this paper, we will show that the rule (3) for probabilities of results of general measurements is a simple consequence of Gleason's theorem. This theorem [6, 7] is a key statement for quantum logics, and also can be considered as a justification of the Born rule [8]. But the usual way of getting the probability rule from Gleason's theorem requires *non-contextuality* to be postulated [1, 8]. We will show that it is possible to avoid the demand of non-contextuality.

We will use Gleason's theorem in the following (somewhat restricted) formulation. Let $p(|e\rangle)$ be a real-valued function of unit vectors $|e\rangle$ in N-dimensional Hilbert space. Suppose that

(1) $N \geq 3$,

(2) the function p is non-negative and continuous,

(3) the value of the sum

$$\sum_{n=1}^{N} p(|e_n\rangle) \tag{4}$$

where unit vectors $|e_1\rangle, |e_2\rangle, ..., |e_N\rangle$ are all mutually orthogonal, does not depend on the choice of the unit vectors.

Then, Gleason's theorem states that the function $p(|e\rangle)$ can be represented as follows:

$$p(|e\rangle) = \langle e| \hat{A} |e\rangle \tag{5}$$

where $\hat{A}$ is some Hermitian operator in the N-dimensional Hilbert space.

We will apply Gleason's theorem in a quite unusual way. Typically, the argument $|e\rangle$ is considered as a property of a measuring device, and the function p as a characteristic of the measured system's state. Our approach is completely reverse—we interpret the vector $|e\rangle$ as a system's state vector, and refer the function p to a measuring device. The main difficulty of this approach consists in satisfying the condition (4). To show that this condition fulfils, we will exploit the concept of environment-induced invariance, or *envariance*, suggested by Zurek [9, 10]. The idea of *envariance* can be formulated as follows: when two quantum systems are entangled, one can *undo* some actions with the first system, performing corresponding counteractions with the second one. Such a possibility of undoing means that these actions do not change the state of the first system (considered as alone) and, in particular, do not change probabilities of results of any measurements on this system [9, 10].

For illustrative purposes, we will depict a quantum system as a moving particle, and a measuring device—as a black box (that emphasizes our ignorance about construction of this device and about processes inside it), see Fig. 1.

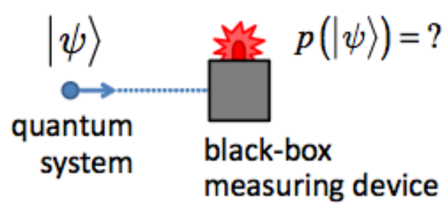

Figure 1: Probability $p(|\psi\rangle)$ of flashing the lamp as a result of interaction of the quantum system in a state $|\psi\rangle$ with the measuring device.

When the particle reaches the black box, the lamp on the box either flashes for a moment, or stays dark. One can introduce the probability $p(|\psi\rangle)$ of flashing the lamp, as the rate of flashes normalized to the rate of particle arrivals, when all these particles are in the state $|\psi\rangle$. The main result of the present paper consists in finding out that

$$p(|\psi\rangle) = \langle\psi|\,\hat{A}\,|\psi\rangle \tag{6}$$

with some Hermitian operator $\hat{A}$.

Though we consider a measurement with only two possible results (flashing and non-flashing of the lamp), this does not lead to any loss of generality. Indeed, one can associate flashing of the lamp with some particular measurement result λ, and non-flashing—with all other results. Then, the function $p(|\psi\rangle)$ in Eq. (6) would be the same as the function $p_\lambda(|\psi\rangle)$ in Eq. (3). So any proof of Eq. (6) also proves Eq. (3), i.e., justifies the POVM nature of every conceivable measurement.

For simplicity, we restrict ourselves by consideration of quantum systems with *finite-dimensional* state spaces.

In Section 2 we will introduce a particular case of en- variance, which will be used later. Section 3 illustrates preparation of a quantum system in a pure state by measurement of *another* system. In Section 4, we will consider a series of thought experiments that combine the features discussed in previous two sections. These experiments show that the function $p(|\psi\rangle)$ obeys Eq. (17). In Section 5 we will demonstrate that Eq. (17) together with Gleason's theorem lead to the probability rule (6). In Section 6, the special case of two-dimensional state space (not covered directly by Gleason's theorem) is considered. Finally, Section 7 shows how the projective postulate (1) (and the Born rule as a particular case) follows from the POVM probability rule (6). Closing remarks are gathered in Section 8.

2. Envariance

Let us consider an experiment shown in Fig. 2.

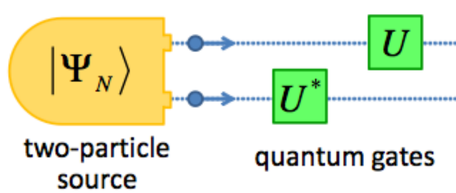

Figure 2: Illustration of *envariance*: the initial joint state of two particles, perturbed by the gate U^*, is restored after applying the gate U to another particle.

Two identical particles are prepared in the joint state

$$|\Psi_N\rangle = \frac{|1\rangle\langle 1| + |2\rangle\langle 2| + ... + |N\rangle\langle N|}{\sqrt{N}} \tag{7}$$

$|1\rangle, |2\rangle, ..., |N\rangle$ being some orthonormal basis of the N-dimensional state space of one particle. After that, each particle passes through a quantum gate, i.e., a device that performs some unitary transformation on the corresponding particle. For the first (upper) particle, an arbitrarily chosen transformation $\hat{U}$ is used. For the second (lower) particle, the complex-conjugated transformation $\hat{U}^*$ (whose matrix elements in the basis $|1\rangle, |2\rangle, ..., |N\rangle$ are complex conjugates to corresponding matrix elements of $\hat{U}$) is applied.

Let us find the joint state $|\Psi'_N\rangle$ of two particles after passing through the gates:

$$|\Psi'_N\rangle = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \left(\hat{U}|n\rangle\right)\left(\hat{U}^*|n\rangle\right) \tag{8}$$

where

$$\hat{U}|n\rangle = \sum_{k=1}^{N} U_{kn}|k\rangle \tag{9}$$

(U_{kn} being matrix elements of $\hat{U}$), and

$$\hat{U}^*|n\rangle = \sum_{l=1}^{N} (U_{ln})^*|l\rangle \tag{10}$$

Substituting the latter two equalities into Eq. (8), and changing the order of summation, one can get

$$|\Psi'_N\rangle = \frac{1}{\sqrt{N}} \sum_{k=1}^{N}\sum_{l=1}^{N} \left(\sum_{n=1}^{N} U_{kn}(U_{ln})^*\right)^* |k\rangle|l\rangle \tag{11}$$

Due to unitarity of the matrix U, the expression in brack??????ets in Eq. (11) reduces to the Kroneker's delta δ_{kl}:

$$\sum_{n=1}^{N} U_{kn}(U_{ln})^* = \delta_{kl} \tag{12}$$

Hence,

$$|\Psi'_N\rangle = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} |k\rangle|k\rangle \equiv |\Psi_N\rangle \tag{13}$$

Thus, effects of two transformations $\hat{U}$ and $\hat{U}^*$, applied to different entangled particles prepared in the state $|\Psi_N\rangle$, Eq. (7), cancel each other. According to Zurek [9, 10], this means that each of these transformations do not change the state of the particle, on which it acts. In other words, the state of each particle is invariant ("*envariant*") under such transformations.

The fact that the two-particle state $|\Psi_N\rangle$ remains unchanged when the particles pass through the gates $\hat{U}$ and $\hat{U}^*$ (Fig. 2) will be used in Section 4.

3. Preparation By Measurement

Let us consider a special measurement device (a "meter") that distinguishes the basis states $|1\rangle, |2\rangle, ..., |N\rangle$ from each other. Therefore the following property is satisfied by definition:

Property a. If a measured system was in the state $|k\rangle$ before measurement by the meter ($k \in \{1, 2, ..., N\}$), then the measurement result will be k with certainty.

It is commonly accepted that any such "meter" has to obey also the following property, which is the reversal of Property a:

Property b. The only state, for which the result of measurement by the meter can be predicted to be k with certainty, is the pure state $|k\rangle$.

Quantum mechanics also guarantees that the following statement is true:

Property c. If two systems were in the joint state $|\Psi_N\rangle$, Eq. (7), and each of them was measured by a meter as shown in Fig. 3, then the results of these two measurements must coincide.

Now consider a state of the upper particle in Fig. 3 just after the lower particle was measured. Let n be the measurement result obtained by the lower meter. Then, according to **Property c**, one can predict that the result of the upper particle's measurement will also be n. Due to **Property b**, this means that the upper particle is now in the pure state $|n\rangle$.

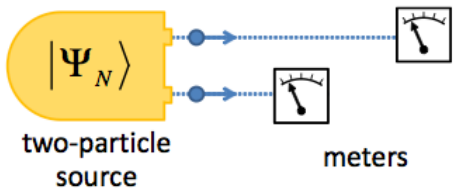

Figure 3: After the measurement of the lower particle, the upper one appears in the state $|n\rangle$, where n is the measurement result.

Hence, if a system of two particles was initially in the state $|\Psi_N\rangle$, and one particle is measured by a meter, this measurement *prepare*s the other particle in the state $|n\rangle$, where n is the result of the measurement. This conclusion will be used in the next Section.

4. Three Thought Experiments

Let us examine the measuring device, schematically represented in Fig. 1, by means of the equipment introduced in Figs. 2 and 3. Figure 4a shows the experiment, in which the source, emitting pairs of particles prepared in the state $|\Psi_N\rangle$, is combined with the measuring device.

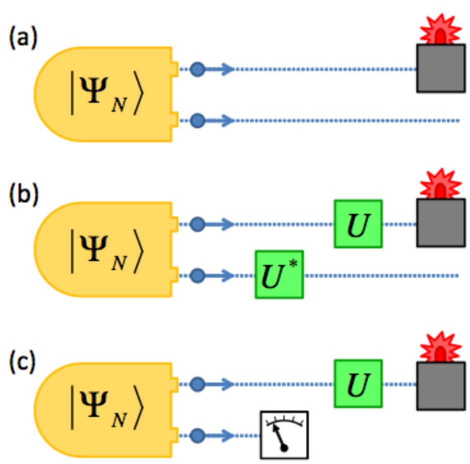

Figure 4: Three thought experiments. The probability $\mathcal{P}$ of flashing the light on the measuring device is the same in all three experiments.

One can define the probability $\mathcal{P}$ of flashing the lamp on the device, as a ratio of the rate of flashing to the rate of emitting the particles by the source.

In the next thought experiment, Fig. 4b, two quantum gates $\hat{U}$ and $\hat{U}^*$, the same as in Fig. 2, are added on the way of particles. It was shown in Section 2 that this combination of gates leaves the state $|\Psi_N\rangle$ unchanged. Thus, from the point of view of the measuring device, nothing was changed when the two gates were introduced, con- sequently the rate of flashing of the lamp remains un- changed. Thus, the probability $\mathcal{P}$ of flashing the lamp on the measuring device is the same in Figs. 4a and 4b.

The third thought experiment in this series (Fig. 4c) differs from the second one (Fig. 4b) by removing the gate U^* and inserting the "meter", which measures the state of the lower particle in the basis $|1\rangle, |2\rangle, ..., |N\rangle$, as in Fig. 3. Since the difference between Fig. 4b and Fig. 4c is related to the lower branch of the experimental setup only, it cannot influence any events of the higher branch. (Otherwise, it would be possible to transfer information from the lower branch to the higher one, without any physical interaction between the branches.) So we conclude that the probability $\mathcal{P}$ of flashing the lamp on the measuring device in the third experiment is the same as in the second one.

Now we will express the value of $\mathcal{P}$ in the third experiment (Fig. 4c) through the function $p(|\psi\rangle)$ defined in Section 1 (a probability of lamp flashing for the pure state $|\psi\rangle$ of the measured particle). Let a_n denote the probability that the meter at the lower branch gives the result N. Also, let $\mathcal{P}_n$ denote the probability that this meter gives the result n *and* the lamp on the measuring device at the higher branch flashes. Obviously,

$$\sum_{n=1}^{N} a_n = 1 \tag{14}$$

$$\sum_{n=1}^{N} \mathcal{P}_n = \mathcal{P} \tag{15}$$

If the lower meter gives the result n, then the upper particle appears in the state $|n\rangle$, according to discussion in Section 3 After passing through the gate U, the upper particle's state turns into $\hat{U}|n\rangle$. Thus, the (conditional) probability of lamp flashing on the device at the higher branch is equal to

$p(\hat{U}\,|n\rangle)$ *if* the lower meter's result is n. Then, according to the multiplicative rule for probabilities,

$$\mathcal{P}_n = a_n p(\hat{U}\,|n\rangle) \tag{16}$$

Combination of Eqs. (15) and (16) gives

$$\sum_{n=1} a_n p(\hat{U}\,|n\rangle) = \mathcal{P} \tag{17}$$

which is simply a manifestation of the law of total probability applied to the experiment shown in Fig. 4c. In Eq. (17), the value of $\mathcal{P}$ does not depend on choice of the unitary operator $\hat{U}$, because this value is the same as in the first experiment (Fig. 4a), see discussion above. Also the values of a_n do not depend on $\hat{U}$.

In the next Section, we will derive Eq. (6) from Eq. (17).

5. Applying Gleason's Theorem

Let $\mathcal{E} = \{|e_1\rangle, |e_2\rangle, ..., |e_N\rangle\}$ be an orthonormal set of vectors in the N-dimensional Hilbert space: $\langle e_m\,|\,e_n\rangle = \delta_{mn}$. Then, it is possible to construct an unitary operator $\hat{U}$ that transforms the set of basis vectors $\{|1\rangle, |2\rangle, ..., |N\rangle\}$ into $\mathcal{E}$:

$$\hat{U}\,|n\rangle = |e_1\rangle\,, \quad n = 1, ..., N \tag{18}$$

Any such unitary operator can be implemented (at least in a thought experiments) as a physical device (quantum gate). Thus, Eq. (17) is valid for the operator $\hat{U}$ defined by Eq. (18). Substituting Eq. (18) into Eq. (17), one can see that

$$\forall\,\mathcal{E} : \sum_{n=1}^{N} a_n p(|e_n\rangle) = \mathcal{P} \tag{19}$$

where values of a_n and $\mathcal{P}$ do not depend on the choice of $\mathcal{E}$.

Now we will see how to get rid of the unknown coefficients a_n. Let us first examine the simplest case of $N = 2$. Eq. (19) for $N = 2$ reads:

$$a_1 p(|e_1\rangle) + a_2 p(|e_2\rangle) = \mathcal{P} \tag{20}$$

If $\{|e_1\rangle, |e_2\rangle\}$ is an orthonormal set, then, obviously, $\{|e_2\rangle, |e_1\rangle\}$ is also an orthonormal set. Therefore Eq. (20) remains valid if one swaps the vectors

$$a_1 p(|e_2\rangle) + a_2 p(|e_1\rangle) = \mathcal{P} \tag{21}$$

Summing up Eqs. (20) and (21), and taking into account that $a_1 + a_2 = 1$, one can arrive to the equality

$$p(|e_1\rangle) + p(|e_2\rangle) = 2\mathcal{P} \tag{22}$$

which is the desired relation between probabilities with- out coefficients a_n

This recipe works also for arbitrary N. Indeed, any permutation of N vectors $|e_n\rangle$ in Eq. (19) gives rise to a valid equality; therefore one can get $N!$ equalities for a given set of vectors. In these $N!$ equalities, each of N vectors $|e_n\rangle$ enters $(N-1)!$ times with each of N factors a_n. Hence, the sum of all these equalities is

$$(N-1)!\left(\sum_{m=1}^{N} a_m\right)\left(\sum_{n=1}^{N} p(|e_n\rangle)\right) = N!\mathcal{P} \tag{23}$$

Finally, taking Eq. (14) into account, one can get the following relation for the function $p(|\psi\rangle)$:

$$\forall\, \mathcal{E} : \sum_{n=1}^{N} p(|e_n\rangle) = N\mathcal{P} \tag{24}$$

One can see now that, for $N \geq 3$, the function $p(|\psi\rangle)$ obeys the conditions of Gleason's theorem. Since p is a probability, it is non-negative. It seems to be physically reasonable to suppose that small variations of the unit vector $|\psi\rangle$ give rise to small variations of the probability $p(|\psi\rangle)$; so we assume that the function p is continuous. (It is possible also to proceed without this continuity assumption, using more sophisticated thought experiments, see Ref. 11.) Finally, the sum (4) is equal to $N\mathcal{P}$ and therefore does not depend on the choice of unit vectors $|e_n\rangle$.

Thus, one can apply Gleason's theorem, that completes the proof of Eq. (6) for the case $N \geq 3$.

6. Case of Two-Dimensional State Space

The above derivation of Eq. (6) does not cover the special case $N = 2$. Now we will see that this case can be reduced to the case $N = 4$.

Consider a system of two non-interacting particles, each of them described by a two-dimensional state space. The first particle is measured by a black-box device, as shown in Fig. 1. As above, we denote as $p(|\psi\rangle)$ the probability of flashing the light on the device, when the state of the *first particle* before its measurement is $|\psi\rangle$. In addition, we denote as $P(|\Psi\rangle)$ the probability of flashing the light, when the joint state of the *two particles* is $|\Psi\rangle$ before measurement of the first particle.

Since $|\Psi\rangle$ is a vector in four-dimensional space ($N = 4$), the above derivation of Eq. (6) is valid for the function $P(|\Psi\rangle)$. Hence, there is such Hermitian operator $\hat{A}$, acting in a four-dimensional space and independent of $|\Psi\rangle$, that

$$P(|\Psi\rangle) = \langle\Psi|\,\hat{A}\,|\Psi\rangle \tag{25}$$

Let us consider the case when the first particle is in some pure state

$$|\psi\rangle = \alpha\,|1\rangle + \beta\,|2\rangle \tag{26}$$

and the second particle is in the state $|1\rangle$. (Here $|1\rangle$ and $|2\rangle$ are some basis vectors in the two-dimensional space.) Then, the joint state of both particles is

$$|\psi\rangle\,|1\rangle \equiv \alpha\,|11\rangle + \beta\,|21\rangle \tag{27}$$

The probability of flashing the light in this situation can be expressed both as $p(|\psi\rangle)$ and $P(|\psi\rangle\,|1\rangle)$

$$p(|\psi\rangle) = P(|\psi\rangle\,|1\rangle) \tag{28}$$

Substituting Eqs. (27) and (25) into Eq. (28), one can express the probability $p(|\psi\rangle)$ as follows:

$$p(|\psi\rangle) = (\alpha^*\,\langle 11| + \beta^*\,\langle 21|)\hat{A}(\alpha\,|11\rangle + \beta\,|21\rangle) \tag{29}$$

i.e.,

$$p(|\psi\rangle) = \begin{pmatrix}\alpha^* & \beta^*\end{pmatrix}\begin{pmatrix}\langle 11|\,\hat{A}\,|11\rangle & \langle 11|\,\hat{A}\,|21\rangle \\ \langle 21|\,\hat{A}\,|11\rangle & \langle 21|\,\hat{A}\,|21\rangle\end{pmatrix}\begin{pmatrix}\alpha \\ \beta\end{pmatrix} \tag{30}$$

The 2×2 matrix in the latter equation can be considered as a representation of some Hermitian operator $\hat{A}_2$, acting in the two-dimensional state space of one particle. Therefore one can rewrite Eq. (30) in the operator form:

$$p(|\psi\rangle) = \langle\psi|\,\hat{A}_2\,|\psi\rangle \tag{31}$$

where $\hat{A}_2$ does not depend on $|\psi\rangle$ (i.e., on α and β).

The derivation of Eq. (31), given in this Section, justifies Eq. (6) for the special case $N = 2$, where N is the dimensionality of the state space of the measured system. Therefore Eq. (6) is now proven for any measurement on any quantum system with finite N.

7. From POVM to the Born Rule

Consider some device that measures an observable $\mathcal{O}$. Let $p(|\psi\rangle)$ be the probability of getting some fixed result λ, when a system in a state $|\psi\rangle$ is measured by this device. It is already proven in Sections 4-6, that the function $p(|\psi\rangle)$ can be represented in the form of Eq. (6), where $\hat{A}$ is some Hermitian operator. In this Section we will see that $\hat{A}$ is a projector onto an eigenspace of the operator $\hat{O}$, which describes the observable $\mathcal{O}$.

Let a matrix A_{mn} represent the operator $\hat{A}$ in a basis $|\varphi_1\rangle, ..., |\varphi_N\rangle$ of eigenvectors of $\hat{O}$:

$$A_{mn} \equiv \langle\varphi_m| \hat{A} |\varphi_n\rangle \tag{32}$$

Then, according to Eqs. (6) and (32), probabilities $p(|\varphi_n\rangle)$ are equal to diagonal matrix elements A_{nn}:

$$p(|\varphi_n\rangle) = \langle\varphi_n| \hat{A} |\varphi_n\rangle = A_{nn} \tag{33}$$

On the other hand, if the state of the measured system is an eigenstate of $\hat{O}$, then the measurement result must be equal to the corresponding eigenvector; therefore $p(|\varphi_n\rangle)$ is 1 if the nth eigenvalue is equal to λ (i.e., if $\hat{O}|\varphi_n\rangle = \lambda|\varphi_n\rangle$), and 0 otherwise. Hence

$$A_{nn} = \begin{cases} 1 & \text{if } \hat{O}|\varphi_n\rangle = \lambda|\varphi_n\rangle \\ 0 & \text{otherwise} \end{cases} \tag{34}$$

Now we will show that non-diagonal matrix elements A_{mn} vanish. For this purpose, let us consider eigenvalues $a_1, ..., a_N$ of the operator $\hat{A}$. Since the trace of a matrix is an invariant, then

$$\sum_n A_{nn} = \sum_k a_k \tag{35}$$

Analogously, since the sum of squared absolute values of all matrix elements is an invariant, then

$$\sum_{m,n} |A_{mn}|^2 = \sum_k a_k^2 \tag{36}$$

Subtracting Eq. (35) from Eq. (36), and taking into account that $|A_{nn}|^2 = A_{nn}$ due to Eq. (34), one can see that

$$\sum_{m \neq n} |A_{mn}|^2 == \sum_k (a_k^2 - a_k) \tag{37}$$

where summation in the left hand side is over all non-diagonal elements.

It is easy to see that all eigenvalues a_k are non-negative. Indeed, if some eigenvalue a_k were negative, then the probability $p(|\chi_k\rangle)$, where $|\chi_k\rangle$ is the corresponding eigenvector, would be negative too:

$$p(|\chi_k\rangle) = \langle\chi_k| \hat{A} |\chi_k\rangle = \langle\chi_k| a_k |\chi_k\rangle = a_k < 0 \tag{38}$$

which is impossible. For a similar reason, a_k cannot be larger than 1. Hence, all eigenvalues a_k are bound within the range [0, 1] and, consequently,

$$\forall n : a_k^2 - a_k \leq 0 \tag{39}$$

Therefore the right hand side of Eq. (37) is negative or zero. But the left hand side of Eq. (37) is positive or zero, so both sides are equal to zero. This proves that all non-diagonal matrix elements A_{mn} vanish.

So the matrix A_{mn} is diagonal, and the action of the operator $\hat{A}$ on basis vectors $|\varphi_n\rangle$ is defined by Eq. (34):

$$\hat{A} |\varphi_n\rangle = A_{nn} |\varphi_n\rangle = \begin{cases} |\varphi_n\rangle & \text{if } \hat{O} |\varphi_n\rangle = \lambda |\varphi_n\rangle \\ 0 & \text{otherwise} \end{cases} \tag{40}$$

The operator $\hat{A}$ is, consequently, the projector onto the eigenspace of $\hat{O}$ with eigenvalue λ. Thus, we have seen that the postulate (1), together with the Born rule (2) in a particular case of non-degenerate eigenvalue λ, are consequences of Eq. (6).

8. Conclusions

In the main part of this paper, Sections 4-6, we have presented a proof that probability of any result of any measurement on a quantum system, as a function of the system?s state vector $|\psi\rangle$, obeys Eq. (6). (For simplicity, only systems with finite-dimensional state spaces were considered.) This justifies the statement that the most general type of measurement in quantum theory is one described by POVM.

It is important to note that this proof of Eq. (6) avoids using the Born rule (or any other form of probabilistic postulate). This opens a possibility of deriving non-circularly the Born rule from Eq. (6). Such possibility is given in Section 7. Note that, despite many efforts aiming to derive the Born rule (see Refs. 8 and 12 for review), there are no generally accepted derivations up to now. Therefore the present approach may be helpful, due to its simplicity: all its essence is contained in three thought experiments shown in Fig. 4.

Finally, let us emphasize the role of entanglement in the present derivation. Though the aim of the derivation was to get a probability rule for *pure* states, the way to such a rule has passed through consideration of the *entangled* state $|\Psi_N\rangle$, Eq. (7). It seems to be that entanglement is a necessary concept for establishing the probabilistic nature of quantum theory.

References

[1] A. Peres, Quantum Theory: Concepts and Methods (Kluwer Academic Publishers, Dordrecht, 1995).

[2] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information, Cambridge Series on Information and the Natural Sciences (Cambridge University Press, 2000).

[3] V. B. Braginsky and F. Y. Khalili, Quantum Measurement (Cambridge University Press, 1995).

[4] C. W. Gardiner, Quantum noise, Springer series in synergetics (Springer-Verlag, 1991).

[5] K. Jacobs and D. A. Steck, Contemporary Physics 47, 279 (2006).

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

[7] A. Dvurečenskij, Gleason's Theorem and Its Applications (Kluwer Academic Publishers, Dordrecht, 1993).

[8] M. Dickson, in Probabilities in Physics, edited by C. Beisbart and S. Hartmann (Oxford University Press, Oxford, 2011) pp. 171-199.

[9] W. H. Zurek, Phys. Rev. Lett. 90, 120404 (2003).

[10] W. H. Zurek, Phys. Rev. A 71, 052105 (2005).

[11] A. V. Nenashev, “Quantum-mechanical measurement apparatus as a black box”, e-print arXiv:1402.2919 [quant-ph].

[12] M. Schlosshauer and A. Fine, Foundations of Physics 35, 197 (2005).