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

A V Nenashev\textsuperscript{1,2}

\textsuperscript{1} Rzhanov Institute of Semiconductor Physics, 630090 Novosibirsk, Russia
\textsuperscript{2} Novosibirsk State University, 630090 Novosibirsk, Russia

E-mail: nenashev@isp.nsc.ru

Received 1 June 2014, revised 14 October 2014
Accepted for publication 27 October 2014
Published 19 December 2014

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

Keywords: Born rule, Gleason’s theorem, envariance

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 $\hat{O}$, 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 the results. Namely, the probability $p_{\lambda}(\psi)$ of the result $\lambda$ is

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

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

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

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

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

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 $\lambda$) 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 equation (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]. However, 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 \( p(\{e\}) \) be a real-valued function of unit vectors \( \{e\} \) in \( N \)-dimensional Hilbert space. Suppose that

1. \( N \geq 3 \),
2. the function \( p \) is non-negative,
3. the value of the sum

\[
\sum_{n=1}^{N} p(\{e_n\}),
\]  

(4)

where unit vectors \( \{|e_1\}, |e_2\}, \ldots|e_N\rangle \) are all mutually orthogonal, does not depend on the choice of the unit vectors.

Then, Gleason’s theorem states that the function \( p(\{e\}) \) can be represented as follows:

\[
p(\{e\}) = \langle e | \hat{A} | e \rangle,
\]  

(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 \( \{e\} \) 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 \( \{e\} \) as a system’s state vector, and refer the function \( p \) to a measuring device. The main difficulty of this approach lies in satisfying the third condition: namely, that the sum (4) is constant. To show that this condition fulfills, we will exploit the concept of environment-induced invariance, or \textit{envariance}, suggested by Zurek [9, 10]. The idea of \textit{envariance} can be formulated as follows: when two quantum systems are entangled, one can \textit{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]. Note that envariance was introduced by Zurek as a tool for understanding the nature of quantum probabilities, and for derivation of the Born rule.

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 figure 1. 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 \( p(\{\psi\}) \) of flashing the lamp, as the rate of flashes normalized to the rate of particle arrivals, when all these particles are in the state \( \{\psi\} \). The main result of the present paper consists of finding out that

\[
p(\{\psi\}) = \langle \psi | \hat{A} | \psi \rangle,
\]  

(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 \( \lambda \), and non-flashing—with all other results. Then, the function \( p(\{\psi\}) \) in equation (6) would be the same as the function \( p_\lambda(\{\psi\}) \) in equation (3). So any proof of equation (6) also proves equation (3), i.e. justifies the POVM nature of every conceivable measurement.

For simplicity, we restrict ourselves to the consideration to quantum systems with \textit{finite-dimensional} state spaces.

In section 2 we will introduce a particular case of \textit{envariance}, 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 the previous two sections. These experiments show that the function \( p(\{\psi\}) \) obeys equation (17). In section 5 we will demonstrate that equation (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. \textbf{Envariance}

Let us consider an experiment shown in figure 2. Two identical particles are prepared in the joint state

\[
|\psi_N\rangle = \frac{|1\rangle|1\rangle + |2\rangle|2\rangle + \ldots + |N\rangle|N\rangle}{\sqrt{N}},
\]  

(7)

\(|1\rangle, |2\rangle, \ldots, |N\rangle\) 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 under the corresponding particle. For the first (upper) particle, an arbitrarily chosen transformation \( \tilde{U} \) is used. For the second (lower) particle, the complex-conjugated transformation \( \tilde{U}^* \) (whose matrix elements in the basis \(|1\rangle, \ldots, |N\rangle\) are complex conjugates to corresponding matrix elements of \( \tilde{U} \)) is applied.

Let us find the joint state \( |\psi'_N\rangle \) of two particles after passing through the gates:

\[
|\psi'_N\rangle = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} (\tilde{U} |\nu\rangle) (\tilde{U}^* |\nu\rangle),
\]  

(8)

where

\[
\tilde{U} |\nu\rangle = \sum_{k=1}^{N} U_{kn} |k\rangle
\]  

(9)

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

\[
\tilde{U}^* |\nu\rangle = \sum_{k=1}^{N} (U_{kn})^* |\nu\rangle.
\]  

(10)

Substituting the latter two equalities into equation (8), and
changing the order of summation, one can get
\[
|\psi_N\rangle = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \sum_{l=1}^{N} \sum_{n=1}^{N} U_{kn}(U_{ln})^* |k\rangle |l\rangle.
\] (11)

Due to unitarity of the matrix \(U\), the expression in brackets in equation (11) reduces to the Kroneker’s delta \(\delta_{kl}:
\[
\sum_{n=1}^{N} U_{kn}(U_{ln})^* = \delta_{kl}.
\] (12)

Hence,
\[
|\psi_N\rangle = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} |k\rangle |k\rangle \equiv |\psi_N\rangle.
\] (13)

Thus, effects of two transformations \(\hat{U}\) and \(\hat{U}^*\), applied to different entangled particles prepared in the joint state \(|\psi_N\rangle\), equation (7), cancel each other. According to Zurek [9, 10], this means that none of these transformations 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 \(U\) and \(U^*\) (figure 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 \(|1\rangle, |2\rangle, \ldots, |N\rangle\) 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 \((k \in \{1, 2, \ldots, N\})\), 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 \(|\psi_N\rangle\), equation (7), and each of them was measured by a meter as shown in figure 3, then the results of these two measurements must coincide.

Now consider a state of the upper particle in figure 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 \(|n\rangle\).

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 prepares 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 figure 1 by means of the equipment introduced in figures 2 and 3. Figure 4(a) shows the experiment, in which the source, emitting pairs of particles prepared in the state \(|\psi_N\rangle\), is combined with the measuring device. One can define the probability \(P\) 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, figure 4(b), two quantum gates \(U\) and \(U^*\), the same as in figure 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; consequently the rate of flashing of the lamp remains unchanged. Thus, the probability \(P\) of flashing the lamp on the measuring device is the same in figures 4(a) and (b).

The third thought experiment in this series (figure 4(c)) differs from the second one (figure 4(b)) by removing the gate \(U^*\) and inserting the ‘meter,’ which measures the state of the
lower particle in the basis \( |1\rangle, |2\rangle, \ldots, |N\rangle \), as in figure 3. Since the difference between figures 4(b) and (c) 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 \( P \) 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 \( P \) in the third experiment (figure 4(c)) through the function \( p(\psi|n) \) 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 \( P_n \) denote the probability that this meter gives the result \( n \) and the lamp on the measuring device at the higher branch flashes. Obviously,

\[
\sum_{n=1}^{N} a_n = 1, \tag{14}
\]

\[
\sum_{n=1}^{N} P_n = P. \tag{15}
\]

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 \( \hat{U} \), 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,

\[
P_n = a_n p(\hat{U}|n\rangle). \tag{16}
\]

A combination of equations (15) and (16) gives

\[
\sum_{n=1}^{N} a_n p(\hat{U}|n\rangle) = P, \tag{17}
\]

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

In the next section, we will derive equation (6) from equation (17).

5. Applying Gleason’s theorem

Let \( \mathcal{E} = \{ |e_1\rangle, |e_2\rangle, \ldots, |e_N\rangle \} \) 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 \( \{ |1\rangle, \ldots, |N\rangle \} \) into \( \mathcal{E} \):

\[
\hat{U}|n\rangle = |e_n\rangle, \quad n = 1, \ldots, N. \tag{18}
\]

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

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

where values of \( a_n \) and \( P \) 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 \( N = 2 \). Equation (19) for \( N = 2 \) reads:

\[
a_1 p(|e_1\rangle) + a_2 p(|e_2\rangle) = P. \tag{20}
\]

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 equation (20) remains valid if one swaps the vectors \( |e_1\rangle \) and \( |e_2\rangle \):

\[
a_1 p(|e_2\rangle) + a_2 p(|e_1\rangle) = P. \tag{21}
\]

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

\[
p(|e_1\rangle) + p(|e_2\rangle) = 2 P, \tag{22}
\]

which is the desired relation between probabilities without coefficients \( a_n \).

This recipe works also for arbitrary \( N \). Indeed, any permutation of \( N \) vectors \( |e_n\rangle \) in equation (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

\[(N - 1)! \left( \sum_{n=1}^{N} a_n \right) \left( \sum_{n=1}^{N} p\left( |e_n\rangle\right) \right) = N! \, P. \]  \hspace{1cm} (23)

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

\[ \forall \, \mathcal{E} : \sum_{n=1}^{N} p\left( |e_n\rangle\right) = N \, P. \]  \hspace{1cm} (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. Finally, the sum (4) is equal to \( N \, P \) and therefore does not depend on the choice of unit vectors \( |e_n\rangle \).

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

### 6. Case of two-dimensional state space

The above derivation of equation (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 figure 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 equation (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

\[ P(|\psi\rangle) = \langle \psi | \hat{A} | \psi \rangle. \]  \hspace{1cm} (25)

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

\[ |\psi\rangle = \alpha |1\rangle + \beta |2\rangle, \]  \hspace{1cm} (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

\[ |\psi\rangle|1\rangle \equiv \alpha |11\rangle + \beta |21\rangle. \]  \hspace{1cm} (27)

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

\[ p(|\psi\rangle) = P(|\psi\rangle|1\rangle). \]  \hspace{1cm} (28)

Substituting equations (27) and (25) into equation (28), one can express the probability \( p(|\psi\rangle) \) as follows:

\[ p(|\psi\rangle) = \left( \alpha^* \langle 11 | + \beta^* \langle 21 | \right) \hat{A} (|11\rangle + |21\rangle), \]  \hspace{1cm} (29)

i.e.

\[ p(|\psi\rangle) = \left( \alpha^* \beta^* \right) \left( \langle 11 | \hat{A} | 11 \rangle \langle 21 | \hat{A} | 21 \rangle \right). \]  \hspace{1cm} (30)

The \( 2 \times 2 \) 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 equation (30) in the operator form:

\[ p(|\psi\rangle) = \langle \psi | \hat{A}_2 | \psi \rangle, \]  \hspace{1cm} (31)

where \( \hat{A}_2 \) does not depend on \( |\psi\rangle \) (i.e. on \( \alpha \) and \( \beta \)).

The derivation of equation (31), given in this section, justifies equation (6) for the special case \( N = 2 \), where \( N \) is the dimensionality of the state space of the measured system. Therefore equation (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 \( \lambda \), when a system in a state \( |\psi\rangle \) is measured by this device. It has already been proven in sections 4–6 that the function \( p(|\psi\rangle) \) can be represented in the form of equation (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 \( |\varphi_1\rangle, \ldots, |\varphi_N\rangle \) of eigenvectors of \( \hat{O} \):

\[ A_{mn} \equiv \langle \varphi_m | \hat{A} | \varphi_n \rangle. \]  \hspace{1cm} (32)

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

\[ p\left( |\varphi_i\rangle \right) = \langle \varphi_i | \hat{A} | \varphi_i \rangle = A_{nn}. \]  \hspace{1cm} (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 \( p(|\varphi_i\rangle) \) is 1 if the rth eigenvalue is equal to \( \lambda \) (i.e. if \( \hat{O}|\varphi_i\rangle = \lambda |\varphi_i\rangle \)), and 0 otherwise. Hence,

\[ A_{nn} = \begin{cases} 1 & \text{if } \hat{O} |\varphi_i\rangle = \lambda |\varphi_i\rangle, \\ 0 & \text{otherwise}. \end{cases} \]  \hspace{1cm} (34)

Now we will show that non-diagonal matrix elements \( A_{mn} \) vanish. For this purpose, let us consider eigenvalues \( a_1, \ldots, a_N \) of the operator \( \hat{A} \). Since the trace of a matrix is an
invariant, then
\[ \sum A_{mn} = \sum a_k. \tag{35} \]

Analogously, since the sum of squared absolute values of all matrix elements is an invariant, then
\[ \sum |A_{mn}|^2 = \sum a_k^2. \tag{36} \]

Subtracting equation (35) from equation (36), and taking into account that $|A_{mn}|^2 = A_{mn}$ due to equation (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 eigenvalues $a_k$ were negative, then the probability $p(\chi_k)$, where $|\chi_k\rangle$ is the corresponding eigenvector, would be negative too:
\[ p(\chi_k) = \langle \chi_k | \hat{A} | \chi_k \rangle = \langle \chi_k | a_k | \chi_k \rangle = a_k < 0, \tag{38} \]
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 equation (37) is negative or zero. However, the left hand side of equation (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 $A_{nn}$ is diagonal, and the action of the operator $\hat{A}$ on basis vectors $|q_k\rangle$ is defined by equation (34):
\[ \hat{A} |q_k\rangle = A_{nn} |q_k\rangle = \begin{cases} |q_k\rangle & \text{if } \hat{O} |q_k\rangle = \lambda |q_k\rangle, \\ 0 & \text{otherwise.} \end{cases} \tag{40} \]

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

8. Conclusions

In the main part of this paper, sections 4–6, we have presented a proof that the probability of any result of any measurement on a quantum system, as a function of the system’s state vector $|\psi\rangle$, obeys equation (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 equation (6) avoids using the Born rule (or any other form of probabilistic postulate). This opens up a possibility of deriving the Born rule non-circularly from equation (6). Such possibility is given in section 7. Note that despite many efforts aiming to derive the Born rule (see [8, 11] for review), there are no generally accepted derivations up to now. Therefore the present approach may be helpful due to its simplicity: its entire essence is contained in three thought experiments shown in figure 4.

Finally, let us emphasize the role of entanglement in the present derivation. Consideration of an entangled state of two particles $|\Psi_N\rangle$, equation (7), has helped us to establish the probability rule for pure states of one particle alone (not entangled with any environment). It seems to be that entanglement is a necessary concept for establishing the probabilistic nature of quantum theory.

References

[1] Peres A 1995 Quantum Theory: Concepts and Methods (Dordrecht: Kluwer)
[2] Nielsen M A and Chuang I L 2000 Quantum Computation and Quantum Information (Cambridge: Cambridge University Press)
[3] Braginsky V B and Khalili F Y 1995 Quantum Measurement (Cambridge: Cambridge University Press)
[4] Gardiner C W 1991 Quantum Noise (Berlin: Springer)
[5] Jacobs K and Steck D A 2006 Contemp. Phys. 47 279
[6] Gleason A M 1957 J. Math. Mech. 6 885
[7] Dvurečenskij A 1993 Gleason’s Theorem and Its Applications (Dordrecht: Kluwer)
[8] Dickson M 2011 Probabilities in Physics ed C Beisbart and S Hartmann (Oxford: Oxford University Press) pp 171–99
[9] Zurek W H 2003 Phys. Rev. Lett. 90 120404
[10] Zurek W H 2005 Phys. Rev. A 71 052105
[11] Schlosshauer M and Fine A 2005 Found. Phys. 35 197