

Article

Quantum Probabilities and Maximum Entropy

Andreas Schlatter

Burghaldeweg 2F, 5024 Küttigen, Switzerland; schlatter.a@bluewin.ch; Tel.: +41-79-870-7775

Received: 22 May 2017; Accepted: 24 June 2017; Published: 26 June 2017

Abstract: Probabilities in quantum physics can be shown to originate from a maximum entropy principle.**Keywords:** quantum measurement; Born-rule; density operator; symmetry

1. Introduction

Ever since quantum physics was discovered at the beginning of the 20th century, there has been a debate about its interpretation. The mathematical formalism does not allow a direct, descriptive interpretation of quantum-objects in space and time and seems to be underdetermined [1] with regard to its ontological structure. As a consequence, there are multiple interpretations/ontologies available today. A particular topic has always been the role and nature of the probabilities in quantum physics. A cornerstone of quantum theory is the fact that the world is empirically and epistemically probabilistic. This means that agents are able to assign probabilities to future events, which are then empirically tested by multiple trials of experiments on identical systems. We also include the term “epistemic”, because of the fact that, although there are today deterministic models of the quantum realm [2], it seems that their values are knowable only modulo randomly distributed initial conditions. The fact that nature shows probabilistic patterns and that agents can theoretically predict and then empirically find them in experiments is by no means self-evident. We will see that once we have defined how physical properties are represented in the theory, we only need two additional, plausible assumptions on how agents empirically gather data and draw conclusions, to uniquely define the theory. The fact that experimental (statistical) frequencies coincide with the probabilities is built into the theory and thence logically no surprise. Again, what is astounding though, is that nature plays the game and allows such a theory in the first place.

There has been a long debate on how to interpret randomness in quantum physics [3]. There are, at first sight, three different kinds of probabilities. The first category consists of the probabilities, which arise from pure quantum states through the Born-rule [4]. These are sometimes considered as the “true” quantum-probabilities. The second category consists of the weights in mixed states and the third one of the frequencies found in multiple trials of an experiment on identical quantum systems. There naturally arises the quest for an underlying principle, common to all categories. Since agents can choose to do a single experiment only, the frequency definition seems to fall short as the common principle. On the other hand and in view of what we said earlier, it also seems a bold standpoint to say that the probabilities are merely subjective [5]. The fact that nature allows a theory, where agents by experiments can test probabilities arising from that theory, does say something about nature itself, of which agents, admittedly, are a part. A further interpretation, which works for the single trial, is the one of propensities [6]. There is a vast literature on the philosophical question of the true nature of probabilities, which we are unable to cover here [7]. We will add a specific unifying view and argue in this paper that the probabilities in quantum theory can be understood as the result of symmetry under permutations in combination with Laplace’s law of indifference, i.e., a maximum entropy principle.

2. Quantum Physics

The ansatz for the mathematical theory of quantum physics is to represent a physical quantity as a self-adjoint operator $A \in L(H)$ in the space of linear operators over a state space H , which carries the structure of a complex Hilbert space of some dimension $d \in \mathbb{N}$. The values, which this quantity can assume in an experiment, are the corresponding eigenvalues $\lambda_k \in \mathbb{R}$, $k \leq d$, of A . We have to find a way to assign probabilities to these eigenvalues, a task which is equivalent to assigning probabilities to the orthogonal projection operators $\{\Pi_k\}_{k \leq d} \in L(H)$, which project states in H to eigenstates corresponding to the λ_k . If agents assign a probability p_0 to a projection operator Π_0 , which is common to two families of orthogonal projectors, $\{\Pi_k\}_{k \leq d}$, $\{\Gamma_k\}_{k \leq d}$, and do corresponding experiments, they would like to be sure that, if they find the frequency p_0 , the results describe the same event Π_0 . This is a non-contextuality condition (note that the set of (conditional) probabilities over realized values is contextual, as a theorem by Kochen–Specker [8] and an example by Hardy [9] show). A famous theorem from Gleason [10] says that to any non-contextual measure μ on the sets of projectors $\{\Pi\}_H$ over a Hilbert space H of dimension $d \geq 3$, there exists a unique positive semi-definite, self-adjoint operator ϱ of trace class one, called density operator, such that $\mu(\Pi) = \text{tr}(\varrho\Pi)$ for all Π . This is the Born-rule. The theorem defines the appropriate measures as well as the quantum states, which are identified with the density-operators ϱ . There is a special class $\mathcal{S}(H)$ of density-operators, called pure states. Every vector, $|\psi\rangle \in H$, defines a corresponding pure state $\varrho = |\psi\rangle\langle\psi|$, which is the projection operator onto $|\psi\rangle$. The set of density operators, $\mathcal{D}(H) \subset L(H)$, is the set of all convex combinations of pure states $\mathcal{S}(H) \subset \mathcal{D}(H)$. The weights corresponding to the convex combinations are the second category of probabilities mentioned above. As important as Gleason's theorem of course is, because it technically defines the right quantum-probabilities, it does not say much about their nature/interpretation.

2.1. Probabilities of Mixed States

The simplest case is the second category, namely the probability weights of mixed states. In the mixed case, there are ex-ante probabilities $p_a \geq 0$, $a \leq M$, $\sum_{a \leq M} p_a = 1$, which generate states of the form

$$\varrho = \sum_{a \leq M} p_a \varrho_a, \quad (1)$$

where the $\{\varrho_a = |\psi_a\rangle\langle\psi_a|\}_{a \leq M}$ are (non-necessarily orthogonal) pure states. These probabilities are considered to be of classical type, i.e., uncertainties about a possible set of preparations. Since the rational numbers $\mathbb{Q}$ are dense in $\mathbb{R}$, we may assume that, with an arbitrarily small error, $p_a = r_a/q_a$, $r, q \in \mathbb{N}_+$. Let $Q = \prod_{a \leq M} q_a$ and $Q_a = \prod_{a' \leq M, a' \neq a} q_{a'}$, respectively. We can then set $N = Q$ and $m_a = Q_a r_a$ to get a number of N states $\{\varrho_{a_k} \varrho_a\}_{a \leq M, k \leq m_a}$ with probabilities $p_{a_k} = 1/N$ and aggregate probabilities $p_a = m_a/N$. This way, the probabilities p_a clearly reflect the indifference principle resulting from the permutation-symmetry of (1).

2.2. Probabilities of Pure States

We now consider a system $\mathcal{S}$ represented by a pure state $|\psi\rangle \in H$ with resolution in the eigenbasis $|e_a\rangle$, $1 \leq a \leq M$, of a self-adjoint operator $A \in L(H)$, $|\psi\rangle = \sum_{a \leq M} \psi_a |e_a\rangle$, $\psi_a \in \mathbb{C}$. We can form the corresponding pure state $\varrho_\psi = |\psi\rangle\langle\psi|$ with matrix-entries $(\varrho_\psi)_{aa'} = \psi_a \psi_{a'}^*$. The Born rule then assigns probabilities

$$p_a = \frac{|\psi_a|^2}{\|\psi\|^2}, \quad (2)$$

to the projectors $\varrho_a = |e_a\rangle\langle e_a|$, $a \leq M$. Assume that there is an additional system $\mathcal{E}$ with orthonormal basis states $\{|n\rangle\}_{n \leq N}$, which is initially in the base state $\varrho_0 = |0\rangle\langle 0|$. A measurement of some state ϱ by the probe $\mathcal{E}$ is an operation U on the joint system $\varrho_{joint} = |0\rangle\langle 0| \otimes \varrho$

$$U(|0\rangle\langle 0| \otimes \varrho)U^*, \quad (3)$$

where U is unitary $UU^* = Id$ (this follows from the fact that a general interaction evolution $U(t) = e^{-(i/\hbar)Ht}$ is unitary). A general unitary transformation on a tensor-product, expressed in the respective bases, can be written as a matrix

$$U = \sum_{an, a'n'} u_{an, a'n'} |a\rangle\langle n| \langle n'| \langle a'| = \sum_{nn'} A_{nn'} \otimes |n\rangle\langle n'|, \quad (4)$$

where the operators $A_{nn'}$ are given by $A_{nn'} = \sum_{aa'} u_{na, n'a'} |a\rangle\langle a'|$. We denote the diagonal sub-block A_{n0} simply by A_n . Since U is unitary, we have

$$\langle 0|UU^*|0\rangle = \sum_n A_n A_n^* = Id. \quad (5)$$

Conversely, we can choose any set of operators A_n satisfying the resolution of the identity-condition (5) to define a measurement on an initial joint state $\varrho_{joint} = |0\rangle\langle 0| \otimes \varrho$. We now have the necessary elements in place to give the main argument.

Assume there is a second system $\mathcal{E}$ with basis $\{|n\rangle\}_{n \leq N}$ and an observer who would like to know in what state $\varrho_a = |e_a\rangle\langle e_a|$ the system $\mathcal{S}$ is in, by making an appropriate measurement U on the joint system $\varrho_{joint} = |0\rangle\langle 0| \otimes \varrho_\psi$. If that is possible in the first place, then, having no additional knowledge, the observer does not, a priori, know in what state $|n\rangle$, $n \leq N$, the probe will be after the measurement and before observation, leading to permutation-symmetry. Let the underlying pure state $|\psi\rangle \in H$ have coefficients $\psi_a = \sqrt{m_a} e^{i\varphi_a}$, $m_a \in \mathbb{N}$, $\varphi_a \in \mathbb{R}$ (since the rational numbers $\mathbb{Q}$ are dense in $\mathbb{R}$, the choice of $m_a \in \mathbb{N}$ is general enough). The probe $\mathcal{E}$ can be chosen appropriately coarse-grained (this coarse-graining is first introduced in [11] in the context of many-worlds) such that $N = \sum_{a \leq M} m_a$. The observer is after the measurement and before observation in a situation where, by lack of further information, she will by Laplace's principle of indifference a priori attribute to each outcome $\langle n|U(|0\rangle\langle 0| \otimes \varrho_\psi)U^*|n\rangle$ equal probability $p_n = 1/N$, $n \leq N$. This attribution is equivalent to maximizing the entropy function $H(p) = -\sum_{n=1}^N p_n \log(p_n)$. The observer can therefore write down in the spirit of (1) an average of outcomes

$$\tilde{q} = \sum_{n \leq N} \frac{1}{N} \left(\langle n|U(|0\rangle\langle 0| \otimes \varrho_\psi)U^*|n\rangle \right) = \sum_{n \leq N} \frac{1}{N} (A_n \varrho_\psi A_n^*). \quad (6)$$

For our purpose, we now chose the operators A_n to be the scaled projectors $\{\tilde{P}_{a_k} (1/\sqrt{m_a}) P_a\}_{a \leq M, k \leq m_a}$ to the basis-states $|e_a\rangle$, $a \leq M$. Note that we have replaced the simple-index n by the double-index a_k . This choice is consistent with the demands of a measurement, since the $\tilde{P}_{a_k}$ satisfy (5)

$$\sum_{n \leq N} \tilde{P}_n^* \tilde{P}_n = \sum_{a \leq M} \sum_{k \leq m_a} \tilde{P}_{a_k}^* \tilde{P}_{a_k} = \sum_{a \leq M} P_a^* P_a = Id. \quad (7)$$

Therefore, we can write (6) in the following form

$$\begin{aligned} \tilde{q} &= \sum_{n \leq N} \frac{1}{N} \left(\tilde{P}_n \varrho_\psi \tilde{P}_n^* \right) = \sum_{a \leq M} \sum_{k \leq m_a} \frac{1}{N} \left(\tilde{P}_{a_k}^* \varrho_\psi \tilde{P}_{a_k} \right) \\ &= \sum_{a \leq M} \frac{1}{N} (P_a \varrho_\psi P_a^*) = \sum_{a \leq M} \frac{m_a}{N} \varrho_a. \end{aligned} \quad (8)$$

Comparing Equation (8) with Equation (1), we see that $\tilde{\varrho}$ can be viewed as a mixed state with probabilities

$$p_a = \frac{m_a}{N} = \frac{|\psi_a|^2}{\|\psi\|^2}, \quad (9)$$

which is the Born-rule.

Before we turn to consider the frequencies, let's have a look at composite systems $\varrho_{12} \in \mathcal{D}(H \otimes H)$. To show the principle it is sufficient to look at binary systems.) The state ϱ_{12} may be mixed or pure and we can apply the findings in a straightforward way. The single components are given by the partial trace $\varrho_{1/2} = \text{tr}_{2/1}(\varrho_{12})$. If the state has the form $\varrho_{12} = \varrho_1 \otimes \varrho_2$, then we are in the separable case and can apply the results in 2.1 and 2.2 to each individual component, which can be pure or mixed. In case ϱ_{12} is entangled, then the partial trace always produces a mixed state. When we now consider frequencies, then the individual quantum systems might be single or composite, what is important is that they are temporally separable in order to allow for statistics.

2.3. Frequencies

The theory so far does only cover single trials. Assume there is a density-operator $\varrho \in \mathcal{D}$ and a complete set of projectors $\{\Pi_k\}_{k \leq M}$. To find probabilities for a sequence of different outcomes $k_1, \dots, k_N$, of N experiments on ϱ (this is done on N identically prepared systems) we can apply Gleason's theorem to the tensor product [5]

$$\varrho^N = \varrho \otimes \dots \otimes \varrho, \quad (10)$$

to get

$$p(k_1, \dots, k_N) = \text{tr}(\varrho^N \Pi_{k_1} \otimes \dots \otimes \Pi_{k_N}) = p_{k_1}, \dots, p_{k_N}, \quad (11)$$

with

$$p_k = \text{tr}(\varrho \Pi_k). \quad (12)$$

So the outcomes of repeated measurements are identically and independently distributed (i.i.d.). The probability for outcome k to occur n_k times, $k \leq M$, $\sum_k n_k = N$, is given by the multinomial distribution

$$p(n_1, \dots, n_M) = (N! / n_1! \dots n_M!) p_1^{n_1} \dots p_M^{n_M}. \quad (13)$$

The individual counting functions n_k are binomially distributed and hence $E(n_k) = N p_k$. For large N the averages, $\bar{n}_k$, of the statistical counting functions approach the expectation values and therefore

$$\bar{n}_k / N \approx p_k. \quad (14)$$

The fact that $\bar{n}_k \rightarrow E(n_k)$, $N \rightarrow \infty$, is due to the law of large numbers. The frequencies with their implied principle of indifference (14) indeed replicate the probabilities. This is achieved by a strong assumption in the theory, reflected in Equations (11), (13), and (14). It is the independence condition for the multi-trial states ϱ^N , $N \in \mathbb{N}$. Actually, it is itself a consequence of the assumption that agents have maximal information about a system of N copies of a quantum state [5]. Independence implies serial permutation-symmetry, i.e., the fact that it does not count in which sequence the results occur. So in the case of multiple-trials the theory uses a stronger assumption than permutation-symmetry to obtain the compatible frequency-probabilities (14). Can we weaken the assumption?

It is remarkable that, due to the (infinite) de-Finetti theorem [12], the assumption of independence can be weakened to the one of exchangeability, to still allow reasonable statistics. Exchangeability stands for permutation-symmetry of the joint distribution of N trials X_n , $n \leq N$,

$$q_N(X_1, \dots, X_N) = q_N(X_{\pi(1)}, \dots, X_{\pi(N)}), \pi \in \text{Per}(N),$$

and for consistency from step N to $N - 1$, $q_{N-1} = \text{tr} q_N$. If satisfied, it can uniquely represent an N -trial state q_N by an integral over product states of form (10) by means of a measure μ on $\mathcal{S}(H)$

$$q_N = \int_{\mathcal{S}(H)} \mu(q) q^N. \quad (15)$$

(The measure μ belongs to the second category of probabilities.) For states q_N of form (15) the statistical approach (14) works with some suitable adjustments (while the distributions are directly integral averages over the product state-distributions, there holds a law of large numbers only conditional to a suitable σ -algebra [13]). Whether we work with states of maximal information (10) or states of form (15), in any case permutation-symmetry and the principle of indifference are key features of the frequencies (14), derived in multiple-trials.

3. Conclusions

We have, in the exposition, not made use of any specific interpretation of quantum mechanics, but relied on the original formalism only. We have seen that it is possible to interpret all the probabilities in quantum physics as arising from permutation-symmetry and the principle of indifference, which amounts to maximum entropy. For the single trial, this simply means that at any single point in time we have a number of equiprobable states which can occur. In the case of multiple trials, any single state can occur equiprobably at a number of different points in time. Since permutation-symmetry is very natural and inherent in statistics, and since Laplace's principle is a basic rational intuition, which underlies elementary combinatorics, we feel that both together are acceptable principles on which to base a theory of nature. So, given the projectors on Hilbert space as the model-structure, the assumptions of non-contextuality and independence/exchangeability lead together with the principle of indifference directly to both a formal and interpretative specification of the probabilities in quantum physics. The thus specified probabilities are real in as much, as they belong to a theory, in which agents and systems enter into a testable relationship. They are hence as much features of agents as they are of the physical systems.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Lewis, P.J. *Phenomena and Theory*; Oxford University Press: Oxford, UK, 2016; pp. 22–24.
2. Bohm, D. A suggested interpretation of the quantum theory in terms of “hidden” variables. *Phys. Rev.* **1952**, *85*, 166–180. [[CrossRef](#)]
3. Lewis, P.J. *Indeterminacy*; Oxford University Press: Oxford, UK, 2016; pp. 72–105.
4. Born, M. Quantenmechanik der Stoßvorgänge. *Z. Phys.* **1926**, *37*, 863–867. (In German) [[CrossRef](#)]
5. Caves, M.C.; Fuchs, C.A.; Schack, R. Quantum probabilities as Bayesian probabilities. *Phys. Rev. A* **2002**, *65*, 022305. [[CrossRef](#)]
6. Popper, K.R. The Propensity Interpretation of Probability. *Br. J. Philos. Sci.* **1959**, *10*, 25–42. [[CrossRef](#)]
7. Childers, T. *Philosophy and Probability*; Oxford University Press: Oxford, UK, 2013.
8. Kochen, S.; Specker, E. The Problem of Hidden Variables in Quantum Mechanics. *J. Math. Mech.* **1967**, *17*, 59–87. [[CrossRef](#)]
9. Mermin, N.D. What is quantum mechanics trying to tell us? *Am. J. Phys.* **1998**, *66*, 753–767. [[CrossRef](#)]

10. Gleason, A.M. Measures on the closed subspaces of a Hilbert space. *J. Math. Mech.* **1957**, *6*, 885–893. [[CrossRef](#)]
11. Zurek, W.H. Environment-assisted invariance, entanglement and probabilities in quantum physics. *Phys. Rev. Lett.* **2003**, *90*, 120404. [[CrossRef](#)] [[PubMed](#)]
12. Caves, M.C.; Fuchs, C.A.; Schack, R. Unknown Quantum States: The Quantum de Finetti Representation. *J. Math. Phys.* **2002**, *49*, 4537–4559. [[CrossRef](#)]
13. Kallenberg, O. *Probabilistic Symmetries and Invariance Principles*, 1st ed.; Springer: New York, NY, USA, 2005; p. 510.



© 2017 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).