Quantum States and Generalized Observables: A Simple Proof of Gleason's Theorem

P. Busch*
Department of Mathematics, University of Hull, Hull HU6 7RX, United Kingdom
(Received 29 May 2003; published 19 September 2003)

A quantum state can be understood in a loose sense as a map that assigns a value to every observable. Formalizing this characterization of states in terms of generalized probability distributions on the set of effects, we obtain a simple proof of the result, analogous to Gleason's theorem, that any quantum state is given by a density operator. As a corollary we obtain a von Neumann–type argument against noncontextual hidden variables. It follows that on an individual interpretation of quantum mechanics the values of effects are appropriately understood as propensities.

DOI: 10.1103/PhysRevLett.91.120403 PACS numbers: 03.65.Ca, 03.65.Ta, 03.67.–a

In this Letter we characterize a notion of quantum states that takes into account the general representation of observables as positive operator valued measurements (POVMs). The idea of a state as an expectation value assignment will be extended to that of a generalized probability measure on the set $\mathcal{E}(\mathcal{H})$ of all effects, that is, the positive operators which can occur in the range of a POVM [1]. All such generalized probability measures are found to be of the standard form, i.e., determined by a density operator. This result constitutes a simplified proof and, at the same time, a more comprehensive variant of Gleason's theorem [2]. The Letter concludes with an application of this result to the question of hidden variables [3].

In the traditional formulation of quantum mechanics in Hilbert space, states are described as density operators and observables are represented as Hermitian operators. Alternatively, and equivalently, experimental events and propositions are represented as orthogonal projection operators, and states are defined as generalized probability measures on the set $\mathcal{P}(\mathcal{H})$ of projections, i.e., as functions $E \mapsto v(E)$ with the properties

(P1) $0 \leq v(E) \leq 1$ for all E;

(P2) $v(I) = 1$;

(P3) $v(E + F + \ldots) = v(E) + v(F) + \ldots$ for any sequence $E, F, \ldots$ with $E + F + \ldots \leq I$.

According to Gleason's theorem [2], all states are given by density operators so that $v(E) = v_\rho(E) = \mathrm{tr}[\rho E]$, provided that the dimension of the complex Hilbert space is at least 3. The duality of states and observables is thus characterized through the trace expression $\mathrm{tr}[\rho E]$, which in the minimal interpretation gives the probability of an outcome associated with E, of a measurement performed on a system in state ρ.

In quantum physics there are many experimental procedures leading to measurements whose outcome probabilities are expectations not of projections but rather of effects. It is therefore natural to define a quantum state as a generalized probability measure not just on $\mathcal{P}(\mathcal{H})$ but on the full set of effects, $\mathcal{E}(\mathcal{H})$, in such a way that the conditions (P1)–(P3) hold for all $E, F, \ldots \in \mathcal{E}(\mathcal{H})$. Note that while for sets of projections the condition $E + F + \ldots \in \mathcal{P}(\mathcal{H})$ is equivalent to $E, F, \ldots$ being mutually orthogonal and thus commuting, the commutativity is no longer necessary for $E + F + \ldots \leq I$ to hold if $E, F, \ldots$ are effects. The following analog of Gleason's theorem then holds.

Theorem. Any generalized probability measure $E \mapsto v(E)$ on $\mathcal{E}(\mathcal{H})$ with the properties (P1)–(P3) is of the form $v(E) = \mathrm{tr}[\rho E]$ for all E, for some density operator ρ.

Proof. It is trivial to see that $v(E) = nv(\frac{1}{n}E)$ for all positive integers. Then it follows immediately that $v(pE) = pv(E)$ for any rational $p \in [0, 1]$. Observe also the additivity and the positivity entail that any generalized probability measure is order preserving, $E \leq F \Rightarrow v(E) \leq v(F)$. Let α be any real number, $0 \leq \alpha \leq 1$. Let p_μ be an increasing sequence of rational numbers in $[0, 1]$ which approaches α from below and q_ν a decreasing sequence of rational numbers in $[0, 1]$ which approaches α from above. It follows that $v(p_\mu E) = p_\mu v(E) \leq v(\alpha E) \leq v(q_\nu E) = q_\nu v(E)$, and therefore, in the limit we have $\alpha v(E) \leq v(\alpha E) \leq \alpha v(E)$. Hence, $v(\alpha E) = \alpha v(E)$.

Let A be any positive bounded operator not in $\mathcal{E}(\mathcal{H})$. We can always write $A = \alpha E$, with $E \in \mathcal{E}(\mathcal{H})$ and suitable $\alpha \geq 1$. [In fact, as A is bounded, there is a positive number α such that $\langle \varphi | A \varphi \rangle \leq \alpha$ for all unit vectors φ. This means that the operator $E := (1/\alpha)A$ is an effect.]

Let $E_1, E_2 \in \mathcal{E}(\mathcal{H})$ be such that $A = \alpha_1 E_1 = \alpha_2 E_2$. Assume without loss of generality that $1 \leq \alpha_1 < \alpha_2$. Then $v(E_2) = \frac{\alpha_1}{\alpha_2} v(E_1)$, and so $\alpha_1 v(E_1) = \alpha_2 v(E_2)$. Thus we can uniquely define $v(A) = \alpha_1 v(E_1)$.

Let A, B be positive bounded operators. Take $\gamma > 1$ such that $\frac{1}{\gamma}(A + B) \in \mathcal{E}(\mathcal{H})$. Then we can write $v(A + B)$ as $\gamma v[\frac{1}{\gamma}(A + B)] = \gamma v(\frac{1}{\gamma}A) + \gamma v(\frac{1}{\gamma}B) = v(A) + v(B)$.

Finally, let C be an arbitrary bounded Hermitian operator. Assume we have two different decompositions $C = A - B = A' - B'$ into a difference of positive operators. We have $v(A) + v(B') = v(B) + v(A')$ and so $v(A) - v(B) = v(A') - v(B')$. Thus we can uniquely define: $v(C) := v(A) - v(B)$. It is now straightforward to verify the linearity of the map v thus extended to the vector space of all bounded linear Hermitian operators. We have found that any generalized probability measure on effects extends to a unique positive linear functional on this vector space. If the Hilbert space $\mathcal{H}$ is finite dimensional, it is a straightforward application of elementary linear algebra to show that any such functional is obtained from a density operator via the trace formula. If the Hilbert space is infinite dimensional, it still holds true that these linear functionals are bounded [already as a consequence of (P1)–(P3) with additivity over finite collections of effects] and even normal (using the additivity over countable sets of effects). In fact, since $C \leq D$ implies $v(C) \leq v(D)$, we have that $|v(C)| \leq \|C\|$ (using $D = \|C\| I$). This is the boundedness of v. (The symbol $\|C\|$ denotes the norm of the bounded Hermitian operator C, defined as the smallest number $\gamma \geq 0$ for which $\|C\varphi\| \leq \gamma \|\varphi\|$ for all $\varphi \in \mathcal{H}$.) To show that v is normal one has to verify that for any norm bounded increasing sequence of Hermitian operators C_n with limit C, it follows that $v(C_n)$ approaches $v(C)$. To see this one uses the following construction to utilize the countable additivity property P3. Take $\alpha > 0$ large enough such that all $C_n + \alpha I \geq O$ and $C + \alpha I \geq O$. Take $\beta > 0$ such that $F = \beta(C + \alpha I) \leq I$. Then the operators $F_n = \beta(C_n + \alpha I)$ and F are effects and we have that the increasing sequence of effects F_n approaches F. Write $E_1 = F_1, E_2 = F_2 - F_1, \ldots, E_n = F_n - F_{n-1}$. These operators E_n are also effects. They satisfy $E_1 + E_2 + \ldots + E_n = F_n$, so that $E_1 + E_2 + \ldots + E_n + \ldots = F$. Therefore $F = \lim_{n\to\infty} F_n = E_1 + E_2 + \ldots + E_n + \ldots$, and by (P3) we obtain $v(F) = \sum v(E_n) = \lim_{n\to\infty} v(F_n)$. This translates into $v(C_n) \to v(C)$. Finally, we note that it is well-known that any positive normal linear functional is obtained from a density operator (e.g., [4], Lemma 1.6.1, or see the direct elementary proof due to von Neumann [5,6]). □

The conclusion of our theorem is the same as that of Gleason's theorem. The extreme simplicity of the proof in comparison to Gleason's proof is due to the fact that the domain of generalized probability measures is substantially enlarged, from the set of projections to that of all effects.

The statement of the present theorem also extends to the case of two-dimensional Hilbert spaces where Gleason's theorem fails. It is worth noting that the dispersion-free generalized probability measures constructed on the set of projections of a two-dimensional Hilbert space (see, e.g., [7,8]), simply do not extend to any generalized probability measures on the full set of effects. The reason must be seen in the fact that the additivity requirement for v on sets of pairwise orthogonal projections is too weak to enforce the linearity of v, considering that such sets of projections are mutually commutative.

Here is a simple intuitive argument demonstrating that there are no linear extensions of any dispersion-free generalized probability measure on the projections of a two-dimensional Hilbert space. We use the Poincaré sphere representation of positive operators of trace 1, $A = 1/2(I + a \cdot \sigma)$, where $\sigma = (\sigma_x, \sigma_y, \sigma_z)$, $a = (a_x, a_y, a_z)$, with $\|a\|^2 = a_x^2 + a_y^2 + a_z^2 \leq 1$. All projections are then either I or O or $P = 1/2(I + n \cdot \sigma)$, with $\|n\| = 1$. Let v be a dispersion-free generalized probability measure on the projections. Any pair of mutually orthogonal projections $P, P' = I - P$ will have values 1 and 0 such that their sum is 1. Hence there are nonorthogonal pairs $P = 1/2(I + n \cdot \sigma)$, $Q = 1/2(I + m \cdot \sigma)$ such that both have value 0. If v had a linear extension, then all the effects corresponding to the line segment joining n and m, $E = \lambda P + (1 - \lambda)Q$, with $0 \leq \lambda \leq 1$, would have values $v(E) = \lambda v(P) + (1 - \lambda)v(Q) = 0$. On the other hand, we can write E in its spectral decomposition $E = \mu R + (1 - \mu)R'$, where $0 < \mu < 1$ if $0 < \lambda < 1$. Assume that $v(R) = 1$, $v(R') = 0$, then $v(E) = \mu \neq 0$, which contradicts the previous conclusion that $v(E) = 0$. Hence there is no consistent linear extension of v.

Up to this point we have restricted ourselves to the minimal interpretation of quantum states and observables, according to which these entities are tools for calculating experimental probabilities. We have shown that, given the set of effects as a representation of all experimental yes-no questions, any quantum state, understood as a generalized probability measure on the set of effects, is given in the familiar way by a density operator.

This result entails a formalization of the well-known fact that quantum mechanics is an irreducibly probabilistic theory: in contrast to classical probability theory, quantum probabilities cannot be decomposed into convex combinations of dispersion-free (that is, $\{0, 1\}$-valued) generalized probability measures.

We conclude with a brief outline of an application of the above result to interpretations of quantum mechanics that go beyond the scope of the minimal interpretation. Such an interpretation will consider observables as representations of properties of a system and effects as yes-no propositions about the possible values of the observables. The role of states will be to assign values to observables and effects. In a deterministic world, one would expect a complete state description to assign one of the values 1 or 0 to each effect of a complete collection E_i (with $\sum E_i = I$), in such a way that 1 occurs exactly once. Thus the sum of the values for all E_i is 1.

This consideration leads to the idea of defining states as effect valuations, that is, as functions $v{:}E \mapsto v(E)$ of effects with the properties $v(E) \geq 0$, and $v(E) + v(F) + \ldots = 1$ if $E + F + \ldots = I$.

It is easy to see that every effect valuation has the properties (P1)–(P3) of generalized probability measures, and conversely. Hence the above theorem entails that any effect valuation is of the form $v(E) = \mathrm{tr}[\rho E]$ for all $E \in \mathcal{E}(\mathcal{H})$ and some density operator ρ.

An interpretation of valuations as truth value assignments would require the numbers $v(E)$ to be either 1 or 0, indicating the occurrence or nonoccurrence of an outcome associated with E. Valuations with this property are referred to as dispersion-free. The above theorem entails immediately that dispersion-free effect valuations which are defined everywhere on $\mathcal{E}(\mathcal{H})$ do not exist. It follows that noncontextual hidden variables, understood as dispersion-free, globally defined, valuations, are excluded in quantum mechanics.

The argument against noncontextual hidden variables thus obtained resembles formally that of von Neumann [5]. However, von Neumann's problematical assumption, that of additivity of a valuation over arbitrary (countable) sets of (commuting or noncommuting) Hermitian operators [9], is here replaced by the requirement of additivity over (countable) sets of effects that add up to I. Such collections of effects constitute a POVM and are thus jointly measurable in a single experiment. It thus makes sense to consider hypothetical simultaneous (hidden, dispersion-free) values of such sets of effects, and hence also the values of sums of effects provided these sums are bounded by I.

In the case of a pure state $\rho = |\varphi\rangle\langle\varphi|$, the occurrence of values $v_\rho(E)$ strictly between 0 and 1 indicates a situation where the property associated with E is objectively indeterminate; that is, its presence or absence is not just subjectively unknown. This interpretation is in accord with the propensity interpretation of probabilities, according to which the number $v_\rho(E)$ gives a measure of the system's objective tendency to trigger an outcome represented by effect E if the state is given by ρ and a measurement is made of a POVM containing E [10].

As an example, E_1 and $E_2 = I - E_1$ could represent the propositions that a quantum particle is in the upper and lower path of an interferometer, respectively. If a pure state ρ is a superposition of states ρ_1, ρ_2 in which E_1 and E_2 are real, respectively [i.e., $v_{\rho_1}(E_1) = v_{\rho_2}(E_2) = 1$], then there is no convex decomposition of that state in terms of valuations which are dispersion-free, even only with respect to E_1, E_2. The fact that $0 < v_\rho(E_i) < 1$ is then an expression of the *indeterminateness* of the properties E_1, E_2 in the state ρ. The most appropriate way of accounting for this situation seems to be to say that the localization of the quantum particle is extended over the space occupied by the two paths of the interferometer. The quantum particle is present, to a degree quantified by the number $v_\rho(E_i)$, in each of the two paths represented by E_i. If forced by a measurement to decide whether to show up in the upper or lower path, it will do so with a propensity quantified by those numbers.

A related interpretation of valuations for unsharp measurements as approximate truth values has recently been advocated by Breuer [11], who applied Gleason's theorem to obtain a Kochen-Specker theorem [8] for unsharp spin observables.

The nonexistence of dispersion-free effect valuations raises the interesting question whether there are subsets in the set of effects, with meaningful structures, on which such dispersion-free valuations can be defined. Interesting constructions demonstrating a positive answer to this question are presented for subsets of projections in [12], or also for effects in [13]. Intuitively, it appears that the valuations of Bub [12] are defined on relatively sparse sets of projections, but these sets do possess some structures that can be argued to be necessary for a consistent set of definite properties; by contrast, the valuations of Kent [13] are defined on "dense" sets of POVMs where it is not obvious that these are equipped with such "logical" structures. The important task remains to explore how far one can go in defining noncontextual dispersion-free valuations on subsets of effects with appropriate structures, without running into conflict with the modified Gleason theorem proven here.

Note added.—An insightful analysis contributing to this aim has recently been made by Appleby [14].

*Electronic address: p.busch@hull.ac.uk

[1] For an introduction into the notions of effects and POVMs and their use in quantum measurements, see, e.g., the monograph of P. Busch, M. Grabowski, and P. Lahti, *Operational Quantum Physics* (Springer, Berlin, 1995), 2nd corrected printing 1997. An "effect" is a Hermitian operator E acting on the Hilbert space $\mathcal{H}$ which is bounded between O and I; that is, $O \leq E \leq I$ (where $A \leq B$ for two operators A, B means $\langle\varphi|A\varphi\rangle \leq \langle\varphi|B\varphi\rangle$ for all vectors $\varphi \in \mathcal{H}$). We are interested only in POVMs corresponding to discrete observables; such a POVM can simply be described as a resolution of the identity operator I, $\sum_n E_n = I$, in terms of effects $E_n \in \mathcal{E}(\mathcal{H})$.

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

[3] An early version of this Letter (P. Busch, e-print quant-ph/9909073) aimed at a "resurrection" von Neumann's infamous no-hidden variables theorem. The simplified proof of Gleason's theorem on which the argument was based attracted considerable interest. For this reason, the focus of the present Letter is on this result concerning the most general quantum mechanical probability measures.

[4] E. B. Davies, *Quantum Theory of Open Systems* (Academic Press, New York, 1976).
[5] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin, 1932).
[6] The statement of the theorem was given without proof in P. Busch, P. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement* (Springer-Verlag, Berlin, 1996), 2nd ed. To the author's knowledge, the credit for this result is due to the late Sławek Bugajski, who envisaged (unpublished) it at a time when the relevance of effects and POVMs for quantum physics was not yet well established. A proof similar to the present one was published in the abstract mathematical context of base norm and order unit spaces by E. Beltrametti and S. Bugajski, J. Math. Phys. (N.Y.) **38**, 3020 (1997).
[7] J. Bell, Rev. Mod. Phys. **38**, 447 (1966).
[8] S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967).
[9] For a review of early criticism of this assumption, cf. M. Jammer, *The Philosophy of Quantum Mechanics* (Wiley, New York, 1974), Sec. 7.4. According to Jammer, the first to raise this point was German philosopher of physics Grete Hermann, Abhandlungen der Fries'schen Schule **6**, 75–152 (1935). For a more recent account, see N. D. Mermin, Rev. Mod. Phys. **65**, 803 (1993).
[10] An interpretation of objective quantum probabilities (in pure states and for projections) in terms of propensities is discussed by N. Gisin, J. Math. Phys. (N.Y.) **25**, 2260 (1984). Here the term "objective" probability refers to the indeterminateness of a property whose probability is not 1 or 0 (in a pure state). Under appropriate conditions one can extend the concept of propensity to mixed states and POVMs, namely, in the case where the effects of the POVM do not commute with the density operator. A propensity interpretation thus understood was envisaged by W. Heisenberg who used terms such as disposition, potentiality, or tendency of actualization. The use of the "propensity" interpretation of probability in the context of quantum mechanics was advocated by K. R. Popper to describe situations of indeterminacy. A dispositional theory of probability was already proposed by C. S. Peirce who used the term "would be." I am indebted to A. Shimony for this piece of information on the history of the propensity interpretation. For references, see M. Jammer, *The Philosophy of Quantum Mechanics* (Ref. [9]), p. 449, footnote 44.
[11] T. Breuer, Phys. Rev. Lett. **88**, 240402 (2002).
[12] J. Bub, *Interpreting the Quantum World* (Cambridge University Press, Cambridge, 1997).
[13] A. Kent, Phys. Rev. Lett. **83**, 3755 (1999).
[14] D. M. Appleby, Phys. Rev. A **65**, 022105 (2002).