Negativity and Contextuality are Equivalent Notions of Nonclassicality

Robert W. Spekkens

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, United Kingdom CB3 0WA

(Received 25 January 2008; published 7 July 2008)

Two notions of nonclassicality that have been investigated intensively are: (i) negativity, that is, the need to posit negative values when representing quantum states by quasiprobability distributions such as the Wigner representation, and (ii) contextuality, that is, the impossibility of a noncontextual hidden variable model of quantum theory. Although both of these notions were meant to characterize the conditions under which a classical explanation cannot be provided, we demonstrate that they prove inadequate to the task and we argue for a particular way of generalizing and revising them. With the refined version of each in hand, it becomes apparent that they are in fact one and the same. We also demonstrate the impossibility of noncontextuality or non-negativity in quantum theory with a novel proof that is symmetric in its treatment of measurements and preparations.

DOI: 10.1103/PhysRevLett.101.020401 PACS numbers: 03.65.Ta, 03.65.Ud

It is common to assert that the discovery of quantum theory overthrew our classical conception of nature. But what, precisely, was overthrown? Being specific about the way in which a quantum universe differs from a classical universe is a notoriously difficult task and continues to be a subject of ongoing research today. This problem has become one of practical concern in quantum information theory and quantum metrology as insights into the differences between the two theories help to identify and analyze information-processing tasks for which quantum protocols outperform their classical counterparts. The two notions of nonclassicality with which we shall be concerned in this article are negativity and contextuality, or more precisely, the presence of negative values in quasiprobability representations of quantum theory [1] and the impossibility of noncontextual hidden variable models [2]. We argue that these notions, construed in the traditional manner, are not sufficiently general and we promote a particular way of generalizing and revising them. In particular, we argue that non-negativity in the distributions representing quantum states is not sufficient for classicality; the conditional probabilities representing measurements must also be non-negative. Furthermore, we argue that a classical explanation cannot be ruled out by considering a single quasiprobability representation, such as the Wigner representation [3]; negativity must be demonstrated to hold for all such representations. Following previous work by the author [4], we also argue that an assumption of determinism that is part of the traditional notion of noncontextuality should be excised and that context independence should be required not just for measurement procedures but for preparation procedures as well. Under these refinements, the two notions of nonclassicality are revealed to be equivalent.

Negativity.—In 1932, Wigner showed that one can represent a quantum state by a function on phase space, now known as the Wigner function, having the property that the marginals over all quadratures (linear combinations of position and momentum) reproduce the statistics for the associated quantum observables [3]. This function cannot, however, be interpreted as a probability distribution over a classical phase space because for some quantum states it is not everywhere non-negative. We shall say that such quantum states exhibit negativity in their Wigner representation. It is commonly thought that such negativity is a good notion of nonclassicality. However, we argue that it is neither a necessary nor a sufficient condition for the failure of a classical explanation.

First, we show that it is not a necessary condition. It is well known that the original Einstein-Podolsky-Rosen two-particle state has a positive Wigner representation [5] so that it can be associated with a classical probability density over the phase space of the two particles (i.e., over local hidden variables). However, it has also been shown that it is possible to violate a Bell inequality with such a state [6]. How can this be? The resolution of the puzzle is that one can only have a classical interpretation of an experiment if both the preparations and the measurements admit a classical interpretation, and in the experiments in question, the measurements that one requires—such as parity measurements—do not admit such an interpretation because the Wigner representations of the projectors have values outside of the interval [0,1] and consequently cannot be interpreted as conditional probabilities. (A similar argument has been made in Ref. [7].)

Neither is the negativity of the Wigner representation sufficient for nonclassicality. For example, if one considers a limited set of preparations and measurements for which the associated density operators and positive operator valued measures (POVMs) are diagonal in some orthogonal basis, then the diagonal components may be interpreted as classical probabilities, yielding a classical explanation of the experimental statistics. Nonetheless, if the diagonalizing basis does not consist of quadrature eigenstates—for instance, if it consists of number eigenstates—then the Wigner representations of these preparations and measurements will not be positive. More generally, negativity of the Wigner representation does not demonstrate that there

 © 2008 The American Physical Society

is not some other representation with respect to which one achieves a classical explanation. Note that what is classical about these explanations is their use of probability theory. We allow the space of physical states over which the probabilities are defined to be arbitrary.

Generalizing the notion of negativity.—The lesson of the above examples is that in evaluating the possibility of a classical explanation of an experiment, one must consider the negativity of not just the representation of preparations but of measurements as well, and one must look at representations other than that of Wigner.

There is a natural class of representations that includes the Wigner representation and that allows one to preserve a notion of nonclassicality as negativity. We call these "quasiprobability representations" and define them by the following features. Every density operator ρ, a positive trace-class operator on a Hilbert space $\mathcal{H}$, is represented by a normalized and real-valued function μ_ρ on a measurable space Λ. That is, $\rho \leftrightarrow \mu_\rho(\lambda)$ where $\mu_\rho\colon \Lambda \to \mathbb{R}$ and $\int \mu_\rho(\lambda)d\lambda = 1$. Similarly, every POVM $\{E_k\}$, a set of positive operators on $\mathcal{H}$ that sum to identity, is represented by a set $\{\xi_{E_k}\}$ of real-valued functions on Λ that sum to the unit function on Λ. That is, $\{E_k\} \leftrightarrow \{\xi_{E_k}(\lambda)\}$, where $\xi_{E_k}\colon \Lambda \to \mathbb{R}$ and $\sum_k \xi_{E_k}(\lambda) = 1$ for all $\lambda \in \Lambda$. [The trivial POVM $\{I\}$ is represented by $\xi_I(\lambda) = 1$.] Finally, the representation must be such that

$$\mathrm{Tr}(\rho E_k) = \int d\lambda \mu_\rho(\lambda) \xi_{E_k}(\lambda). \tag{1}$$

There are infinitely many such representations one could define, but popular alternatives to Wigner include the Q and P representations of quantum optics.

We define a *non-negative* quasiprobability representation of quantum theory as one for which

$$\mu_\rho(\lambda) \geq 0, \qquad \xi_E(\lambda) \geq 0 \tag{2}$$

for all density operators ρ and all positive operators E less than identity (i.e., all possible POVM elements). This would constitute a classical representation of all possible preparations and measurements.

Contextuality.—The traditional notion of a noncontextual hidden variable model of quantum theory can be expressed as follows [4]. Denoting a complete set of variables in the model by λ, and the measurable space of these by Λ, one represents every pure quantum state $|\psi\rangle$ by a normalized probability density on Λ, $\mu_\psi(\lambda)$, and every projector-valued measure $\{\Pi_k\}$ (the spectral elements of a Hermitian operator) by a set $\{\xi_{\Pi_k}(\lambda)\}$ of $\{0, 1\}$-valued indicator functions on Λ. An indicator function $\xi_{\Pi_k}(\lambda)$ specifies the probability of outcome k given λ. Because *some* outcome must occur, indicator functions associated with a complete set of outcomes sum to 1, i.e., $\sum_k \xi_{\Pi_k}(\lambda) = 1$ for all λ. A $\{0, 1\}$-valued indicator function is one for which $\xi_{\Pi_k}(\lambda) \in \{0, 1\}$, so that the outcome of the measurement is determined by λ (rather than being probabilistic). We refer to this restriction on indicator functions as the assumption of outcome determinism for sharp measurements. Note that this assumption is part of the traditional notion of noncontextuality (a point to which we shall return). Finally, in order for the hidden variable model to reproduce the probability of outcome k given ψ, one requires that $\int d\lambda \mu_\psi(\lambda) \xi_{\Pi_k}(\lambda) = \langle \psi | \Pi_k | \psi \rangle$.

To see why such a model is called "noncontextual", note that whenever one of the Π_k has rank two or greater, it can be decomposed into a sum of smaller rank projectors in many different ways, and each of these corresponds to a different way of implementing the measurement—a different context. The representation of the measurement in the hidden variable model is presumed to depend only on the Π_k, and not on how the measurement was implemented. The representation is therefore independent of the context, hence noncontextual. The Bell-Kochen-Specker theorem establishes that such a representation of quantum theory is impossible [2].

Generalizing the notion of contextuality.—As argued in detail in previous work [4], the issue of whether a measurement's representation in the model is context-dependent or not can and should be separated from the issue of whether the outcome of the measurement is determined uniquely or only probabilistically by λ. Thus, whereas traditionally the question of interest is whether or not the measurement outcome for a given λ depends on the context of the measurement, we claim that the interesting question is whether the *probabilities* of different outcomes for a given λ depend on the context. This is analogous to Bell's generalization of the notion of locality from measurement outcomes being causally independent of parameter settings at spacelike separation to the probabilities of measurement outcomes being so. Mathematically, the proposed generalization corresponds to dropping the assumption of outcome determinism for sharp measurements from the definition of noncontextuality (by not requiring that $\xi_{\Pi_k}(\lambda)$ be $\{0, 1\}$-valued).

The fully general definition of a noncontextual hidden variable model requires that all procedures—preparations, transformations, and both projective and nonprojective measurements—are represented in a manner that depends only on how the procedure is represented in the quantum formalism. If two procedures differ in ways that are not reflected in the quantum formalism we call this difference part of the context of the procedure.

To be specific, the assumption of preparation noncontextuality is that the probability distribution $\mu_P(\lambda)$ associated with a preparation procedure P depends only on the density operator ρ associated with P, i.e., $\mu_P(\lambda) = \mu_\rho(\lambda)$. For instance, if the preparation is a mixture of pure states $|\psi_k\rangle$ with weights w_k, then the distribution depends only on the average density operator $\rho = \sum_k w_k |\psi_k\rangle\langle\psi_k|$ and not the particular ensemble. Similarly, the assumption of measurement noncontextuality is that the indicator function $\xi_{M,k}(\lambda)$ representing outcome k of a measurement procedure M depends only on the associated POVM element E_k, i.e., $\xi_{M,k}(\lambda) = \xi_{E_k}(\lambda)$.

In order to highlight the content of the assumption of noncontextuality, it is useful to formalize the assumptions that define a hidden variable model of quantum theory. In fact, the set of models that we characterize includes the case wherein the quantum state is a complete description of reality and so it is better to refer to these simply as "ontological models". Every preparation procedure P that is permitted by the theory is represented by a normalized and positive function on a measurable space Λ, and every measurement procedure M is represented by a set of positive functions on Λ that sum to unity. Specifically, we have $P \leftrightarrow \mu_P(\lambda)$, where $\mu_P\colon \Lambda \rightarrow \mathbb{R}$ such that $\int_\Lambda \mu_P(\lambda) d\lambda = 1$, and $M \leftrightarrow \{\xi_{M,k}(\lambda)\}$, where $\xi_{M,k}\colon \Lambda \rightarrow \mathbb{R}$ such that $\sum_k \xi_{M,k}(\lambda) = 1$ for all $\lambda \in \Lambda$, and

$$\mu_P(\lambda) \geq 0, \qquad \xi_{M,k}(\lambda) \geq 0. \tag{3}$$

Finally, let $\mathcal{P}_\rho$ denote all preparation procedures consistent with the density operator ρ, and $\mathcal{M}_{\{E_k\}}$ all measurement procedures consistent with the POVM $\{E_k\}$. An ontological model of quantum theory is such that for all $P \in \mathcal{P}_\rho$, and for all $M \in \mathcal{M}_{\{E_k\}}$,

$$\int d\lambda \mu_P(\lambda) \xi_{M,k}(\lambda) = \mathrm{Tr}(\rho E_k). \tag{4}$$

A *noncontextual* ontological model of quantum theory (in our generalized sense) is one that satisfies

$$\mu_P(\lambda) = \mu_\rho(\lambda) \quad \text{for all } P \in \mathcal{P}_\rho \tag{5}$$

$$\xi_{M,k}(\lambda) = \xi_{E_k}(\lambda) \quad \text{for all } M \in \mathcal{M}_{\{E_k\}}. \tag{6}$$

Equations (5) and (6) codify the assumptions of noncontextuality for preparations and measurements, respectively.

Whatever reasons one can provide in favor of the assumption of measurement noncontextuality [for instance, that it is the simplest possible explanation of the context independence of the right-hand side of Eq. (4)], the very same reasons can be given in favor of the assumption of preparation noncontextuality. Thus if one takes noncontextuality for measurements as a condition for classicality, then noncontextuality for preparations should also be required.

Equivalence of the two notions of nonclassicality.—By substituting the conditions for preparation and measurement noncontextuality, Eqs. (5) and (6), into the conditions for an ontological model, Eqs. (3) and (4), we obtain the conditions for a non-negative quasiprobability representation of quantum theory, Eqs. (1) and (2). So we see that by these definitions, a noncontextual ontological model of quantum theory exists if and only if a non-negative quasiprobability representation of quantum theory exists.

What we have discovered by this analysis is that the assumption of noncontextuality (in our generalized sense) has always been implicit in the notion of a quasiprobability representation. Given its conceptual significance and mathematical simplicity, it is surprising that this connection has not been noted previously. Two likely reasons for this are: (i) the lack of emphasis on the representation of measurements in discussions of negativity, and (ii) the lack of a generalization of contextuality to preparations and nonprojective measurements and the failure to distinguish the assumption of measurement noncontextuality from that of outcome determinism.

No-go theorems for non-negativity or noncontextuality.—An ontological model of quantum theory that is noncontextual, in the generalized sense described here, is impossible [4]. It follows that a non-negative quasiprobability representation of quantum theory is also impossible. This fact is unlikely to surprise those who know quantum theory well. Nonetheless, to our knowledge, it has not been demonstrated previously (although Montina [8] came close to doing so, as we discuss in the conclusions).

An unfortunate feature of existing no-go theorems for noncontextual models is that they do not proceed directly from the assumption of generalized noncontextuality to a contradiction. For instance, in Ref. [4], it is shown that one can base such a proof on the Bell-Kochen-Specker theorem; however, the contradiction is derived not only from the assumption of noncontextuality for sharp measurements, but also the assumption of outcome determinism for sharp measurements, and the latter assumption is in turn derived from noncontextuality for preparations. So, despite the standard impression that these no-go theorems concern only the representation of measurements, we see that the representation of preparations enters the analysis in an indirect way. Similarly, in no-go theorems that appeal only to the assumption of noncontextuality for preparations ([4], Sec. IV), one still relies on the fact that only preparations associated with probability distributions that are nonoverlapping can be discriminated by a single-shot measurement. Thus the representation of measurements has appeared, in an indirect way, within a proof based primarily on the representation of preparations. A proof that is even handed in its treatment of preparations and measurements would be preferable and we now provide one.

An even-handed no-go theorem for non-negativity or noncontextuality.—Suppose that one has a set of preparation procedures associated with density operators ρ_j. A procedure associated with the mixture $\rho \equiv \sum_j w_j \rho_j$ can be implemented as follows. Sample an integer j from the probability distribution w_j and implement the preparation procedure associated with ρ_j. If the distributions over λ that represent each of these procedures in the ontological model are denoted by $\mu_{\rho_j}(\lambda)$, then clearly the distribution that represents the mixture ρ is $\mu_\rho(\lambda) = \sum_j w_j \mu_{\rho_j}(\lambda)$. Thus in a noncontextual ontological model,

$$\text{if } \rho = \sum_j w_j \rho_j, \quad \text{then } \mu_\rho(\lambda) = \sum_j w_j \mu_{\rho_j}(\lambda). \tag{7}$$

A similar argument concerning a mixture of measurements, each of which has a distinguished outcome associated with a positive operator E_j, establishes that in a noncontextual ontological model,

$$\text{if } E = \sum_j w_j E_j \quad \text{then } \xi_E(\lambda) = \sum_j w_j \xi_{E_j}(\lambda). \tag{8}$$

A function f on the space $\mathcal{L}(\mathcal{H})$ of linear operators on $\mathcal{H}$ is convex linear on a convex set $\mathcal{S} \subset \mathcal{L}(\mathcal{H})$ if $f(\sum_k w_k R_k) = \sum_k w_k f(R_k)$, for $R_k \in \mathcal{S}$ where w_k is a probability distribution. Equations (7) and (8) assert that μ as a function of ρ is convex linear on the convex set of density operators, and ξ as a function of E is convex linear on the convex set of positive operators less than identity.

The first two steps of the no-go theorem are familiar as key elements of the generalization of Gleason's theorem [9] to POVMs [10,11] and analogous reasoning plays an important role in Hardy's axiomatization of quantum theory [12]. The first step is to note that a function f that is convex linear on a convex set $\mathcal{S}$ of operators that span the space of Hermitian operators (and that takes value zero on the zero operator if the latter is in $\mathcal{S}$) can be uniquely extended to a linear function on this space. Specifically, if A is a Hermitian operator that can be decomposed as $A = \sum_k a_k R_k$, where $R_k \in \mathcal{S}$, then the extension is $f(A) = \sum_k a_k f(R_k)$.

The second step is to note that by Reisz's representation theorem, the linear function $f(A)$ can be written as the Hilbert-Schmidt inner product of A with some fixed Hermitian operator, say $B^\dagger$, so that $f(A) = \mathrm{Tr}(AB)$.

It follows that

$$\mu_\rho(\lambda) = \mathrm{Tr}[\rho F(\lambda)], \qquad \xi_E(\lambda) = \mathrm{Tr}[\sigma(\lambda) E], \tag{9}$$

where F and σ are functions from Λ to the Hermitian operators on $\mathcal{H}$.

A noncontextual ontological model, or equivalently, a non-negative quasiprobability representation, is one for which $\mu_\rho(\lambda) \geq 0$ and $\xi_E(\lambda) \geq 0$ for all ρ and E, which implies that the operators $F(\lambda)$ and $\sigma(\lambda)$ are not merely Hermitian but positive as well. Given that $\int \mu_\rho(\lambda) d\lambda = 1$, it follows that $\int F(\lambda) d\lambda = I$, and so we can conclude that $F(\lambda) d\lambda$ is a POVM. Furthermore, given that $\xi_I(\lambda) = 1$ for all λ, we also have $\mathrm{Tr}[\sigma(\lambda)] = 1$ and so we can conclude that σ is a map from Λ to density operators.

We now show, using a proof by contradiction, that $\mu_\rho(\lambda)$ and $\xi_E(\lambda)$ of this sort cannot reproduce the quantum predictions. To do so, $\mu_\rho(\lambda)$ and $\xi_E(\lambda)$ would need to satisfy Eq. (1), which implies, via Eq. (9), that $\int d\lambda \mathrm{Tr}[\rho F(\lambda)] \mathrm{Tr}[\sigma(\lambda) E] = \mathrm{Tr}(\rho E)$ for all ρ and E. Given that the set of density operators spans the operator space $\mathcal{L}(\mathcal{H})$, we can infer that $E = \int d\lambda \xi_E(\lambda) F(\lambda)$; i.e., every E is a positive combination of the $F(\lambda)$. Now consider a POVM $\{E_k\}$ with rank-1 elements. Any given element E_k is a positive combination of the $F(\lambda)$, specifically, $E_k = \int d\lambda \xi_{E_k}(\lambda) F(\lambda)$. However, a rank-1 positive operator admits only trivial positive decompositions into positive operators (namely, into ones that are proportional to itself). It follows that $F(\lambda) \propto E_k$ for all λ in the support of ξ_{E_k}. Recalling that for every $\lambda \in \Lambda$, there exists a k such that λ is in the support of ξ_{E_k}, it follows that for every $\lambda \in \Lambda$, there exists a k such that $F(\lambda) \propto E_k$. Repeating the argument for another POVM with rank-1 elements, say $\{E'_j\}$, we conclude that for every $\lambda \in \Lambda$, there exists a j such that $F(\lambda) \propto E'_j$. However, given that no element of $\{E_k\}$ needs to be proportional to any element of $\{E'_j\}$ (for instance, they may be the projector-valued measures corresponding to two bases having no elements in common), we arrive at a contradiction.. [Noting from Eqs. (1) and (9) that $\{F(\lambda)\}$ and $\{\sigma(\lambda)\}$ are dual frames in the operator space, this result implies that the dual of a frame of positive operators cannot also be a frame of positive operators. A direct proof of this fact is possible [13] and provides a faster route to the contradiction.]

A similar argument to the one just provided can be found in the recent work of Montina [8], where it is demonstrated that to avoid negative probabilities in ontological models, the representation of pure states cannot depend bilinearly on the wave function. Although the representation of mixed quantum states is not discussed, it is a short step from this result to a demonstration of a failure of preparation noncontextuality [14].

Attempts to characterize nonclassicality from either the perspective of hidden variables or that of quasiprobability representations drive one to the same conclusion: that the only way in which one can salvage the possibility of an ontological model is to deny the implicit starting point of these representations, the assumption of noncontextuality.

The author gratefully acknowledges J. Barrett, E. Galvão, L. Hardy, M. Leifer, A. Montina, T. Rudolph, and J. Sipe for discussions on these issues.

[1] N. Lütkenhaus and S. M. Barnett, Phys. Rev. A **51**, 3340 (1995).
[2] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966); S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967).
[3] E. P. Wigner, Phys. Rev. **40**, 749 (1932).
[4] R. W. Spekkens, Phys. Rev. A **71**, 052108 (2005).
[5] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, England, 1987), Chap. 21.
[6] K. Banaszek and K. Wodkiewicz, Phys. Rev. A **58**, 4345 (1998).
[7] M. Revzen *et al.*, Phys. Rev. A **71**, 022103 (2005).
[8] A. Montina, Phys. Rev. Lett. **97**, 180401 (2006).
[9] A. M. Gleason, J. Math. Mech. **6**, 885 (1957).
[10] P. Busch, Phys. Rev. Lett. **91**, 120403 (2003).
[11] C. M. Caves, C. A. Fuchs, K. Manne, and J. M. Renes, Found. Phys. **34**, 193 (2004).
[12] L. Hardy, arXiv:quant-ph/0101012.
[13] C. Ferrie and J. V. Emerson, arXiv:0711.2658v2.
[14] The details of this argument can be found in R. W. Spekkens, arXiv:0710.5549.