Available online at www.sciencedirect.com
ScienceDirect
Physics Letters A 365 (2007) 268–272
PHYSICS LETTERS A
www.elsevier.com/locate/pla

POVMs and hidden variables

Allen Stairs

Department of Philosophy, University of Maryland, College Park, MD, USA

Received 22 June 2006; received in revised form 5 January 2007; accepted 15 January 2007

Available online 23 January 2007

Communicated by P.R. Holland

Abstract

Recent results by Paul Busch and Adán Cabello claim to show that by appealing to POVMs, non-contextual hidden variables can be ruled out in two dimensions. While the results of Busch and Cabello are mathematically correct, interpretive problems render them problematic as no hidden variable proofs.

PACS: 03.65.Ca; 03.65.Ta; 03.67.-a

1. Introduction

It is well known that so-called "non-contextual hidden variables" are mathematically possible for the spin observables of a qubit. This raises the question of whether a qubit is fully a quantum system [1]. Several recent results suggest that if we ask the question in the right way, the answer will turn out to be yes. In particular, Paul Busch [2] has proved a version of Gleason's theorem [3] that holds in two dimensions, and Adán Cabello [4] has proved a combinatorial no-hidden variables result in the style of Kochen and Specker [5], but also in two dimensions. What makes these developments possible is the use of POVMs. The new results are welcome, but their interpretation is problematic for two reasons. The first is that, as we will see, there is a straightforward classical model for Cabello's POVMs. The second is a more general problem of specifying which POVMs it would be reasonable to expect the hidden variable theorist to provide values for in the first place.

2. Hidden variables

The term "hidden variable theory" has been used to for a variety of schemes. The sort that bear on the results of Busch and Cabello assume that there is a unique physical quantity corresponding to each self-adjoint operator, and that these quantities have determinate values, revealed by ideal measurement. However, theories that posit determinate values come in different kinds. At one extreme is Kochen and Specker's so-called "trivial hidden variable theory", which ignores all the algebraic constraints that Hilbert space imposes on quantities. At the other extreme is the kind of theory that von Neumann's proof [6] would rule out: one that transfers the additivity of quantum mechanical expected values to the hypothetical dispersion-free states. While this requirement may seem superficially plausible, Bell [7] argued that this impression evaporates on close inspection. If A and B do not commute, then in general there is no way to measure A and B separately and compare the sum of the observed values to the measured value of $A + B$.

So-called "non-contextual deterministic hidden variables" fall between these extremes. If Λ is the set of "hidden states" and A is a measurable quantity, then the value of A is the real number $a = \lambda(A)$, where $\lambda \in \Lambda$, and is the value that an ideal measurement of A would reveal. Associated with each quantum state is a measure on the measurable subsets of Λ, providing marginals that reproduce the quantum probabilities. If such hidden variables were possible, the measurement problem would be dissolved. But it would be dissolved in a way that respects important features of the structure of Hilbert space. Each self-adjoint operator would represent a single quantity whose value is independent of the algebraic context: if $A = f(B) = g(C)$, where the pairs A, B and A, C commute, the values would sat-

E-mail address: stairs@umd.edu.

0375-9601/$ – see front matter
doi:10.1016/j.physleta.2007.01.019

isfy $a = f(b) = g(c)$. This non-contextualism would also have operational meaning: the result of measuring A would be the same whether we measured A directly or instead measured either of B or C and computed the value of A. (See [5] for further detail.)

Familiar results such as Gleason's and Kochen and Specker's rule this sort of hidden variable scheme out so long as the Hilbert space is at least three-dimensional. However, the qubit escapes these proofs. Suppose σ_a and σ_b are spin observables for non-parallel directions a and b. Then σ_a and σ_b have no eigenspaces in common, and the only self-adjoint operators that are functions of both are multiples of the identity. The consequence is that σ_a and σ_b can be represented by algebraically independent subalgebras of a classical probability space, and probabilities can be assigned to them independently. Put another way, Kochen and Specker's trivial hidden variable construction can be used to represent the spin quantities of the qubit without any violation of their strictures on functional relations.

Other sorts of constructions, apparently less trivial than Kochen and Specker's trivial hidden variable theory, are possible as well. Bell [7] provides one. So do Kochen and Specker themselves, and it is worth looking more closely at the details. We will review a sketch here. The reader may consult Section 6 of [5] for details, particularly on the way that probabilities are handled.

The Kochen and Specker model associates the spin observables σ_x, σ_y and σ_z with an orthonormal basis $\{i, j, k\}$ in E^3. Other spin observables are paired with vectors in E^3 in the obvious way. The phase space is the surface of the unit sphere S^2. Each spin operator σ_q picks out a point P_q on S^2—the point at the tip of the vector $q_1 i + q_2 j + q_3 k$. Denote the half-sphere centered at P_q by S_q^+. If the quantum state ψ is the spin-up eigenstate of σ_α, we pair this state with the point P_α on S^2. The hidden state is associated with another point T. Assume that there is a unit disc D, tangent to S^2, with its center at P_α. A point particle is placed on D and "shaken" randomly—i.e., so that the probability for the particle to end up in a given region is proportional to the area of the region. The point T is determined by dropping a perpendicular from the particle to the surface of S^2. The outcome of a measurement of the spin quantity σ_d takes the value $+1$ if T is in the half-sphere S_d^+. Otherwise, the result is -1. Thus, one aspect of the model is "deterministic"; given the hidden state T, the outcome of any measurement is fixed. But the model is *not* deterministic all things considered. The random "shaking" of D takes place anew with each measurement. The measurement result picks out a new quantum state, associated with a new point $S_{\alpha'}^+$ on S^2, and it also picks out a new hidden state. The disc D is moved to the point $S_{\alpha'}^+$ and shaken as before to determine a new hidden state T'. Measurement induces a random change in the hidden variable; the change in the total state of the particle is indeterministic.

Whether or not the sort of indeterminism displayed by this model is inevitable, it certainly is not unreasonable. As we have noted, any NCD HV will have to account for measurement disturbance, and even if the hidden variables assign determinate values to quantities, there is no good a priori reason for insisting that the measurement disturbance must be deterministic. Better, then, to talk of non-contextual *determinate* hidden variables—for short, NCD HVs.

3. Busch's result

Let $P(H)$ be the set of all projectors on a Hilbert space H. Busch points out that we can define a generalized probability measure ν on $P(H)$ as follows:

(1) $0 \leqslant \nu(P) \leqslant 1$,
(2) $\nu(I) = 1$,
(3) $\nu(P_1 + P_2 + \cdots) = \nu(P_1) + \nu(P_2) + \cdots$ for $P_1 + P_2 + \cdots \leqslant I$.

We can think of the quantum state as the measure ν, and if H is at least three-dimensional, Gleason's theorem tells us that given a state ν, there is a density operator ρ such that for each P, $\nu(P) = \mathrm{tr}[\rho P]$. However, in the two-dimensional case there are dispersion-free measures satisfying (1)–(3), even though no density operator captures such measures.

Busch proposes generalizing the notion of "measure". He writes:

> In quantum mechanics, there are many experimental procedures leading to measurements whose outcome probabilities are expectations not of projections but rather of effects [2, p. 120403-1].

"Effects" are positive self-adjoint operators of trace less than or equal to one. A positive operator-valued measure, or POVM, is a set of effects that sums to I, and the effects need not commute. Generalized measurements are represented by POVMs, and each outcome of a generalized measurement corresponds to an effect. We can therefore generalize the notion of a probability measure in (1)–(3) so that the measure is defined on the set $E(H)$ of all *effects* on H, and satisfies the conditions

(1′) $0 \leqslant \nu(E) \leqslant 1$,
(2′) $\nu(I) = 1$,
(3′) $\nu(E_1 + E_2 + \cdots) = \nu(E_1) + \nu(E_2) + \cdots$ for $E_1 + E_2 + \cdots \leqslant I$.

As Busch shows, each ν has a unique linear extension to all bounded self-adjoint operators on H. Straightforward considerations entail that ν, thus extended, picks out a unique density operator ρ such that for any bounded self-adjoint operator A, $\nu(A) = \mathrm{tr}(\rho A)$. In particular, the value $\nu(E)$, where E is an effect, must be $\mathrm{tr}(\rho E)$. This applies, of course, even if the effect E is a projector, and the theorem holds even in two dimensions.

The effects in the sum $E_1 + E_2 + \cdots$ need not commute, but the usual objection to von Neumann's argument does not apply. If $\{E_1, E_2, \ldots\}$ is a POVM, then there will be an experiment whose outcomes are associated with the E_i—it will be possible to "measure" the E_i simultaneously.

This is an elegant result. Busch notes that, interpreted minimally, it tells us that if we use effects on Hilbert space to represent experimental yes–no questions, then the usual way of computing probabilities via the trace rule is the only possible way. However, he suggests that the result can be interpreted more robustly: as a proof that there are no non-contextual hidden variables. He writes:

> The above theorem entails immediately that dispersion-free effect valuations which are defined everywhere on $E(H)$ do not exist. It follows that non-contextual hidden variables, understood as dispersion-free, globally defined, valuations which are defined everywhere on $E(H)$ do not exist [2, p. 120403-3].

As a purely mathematical claim, this is certainly correct. Whether it should trouble the NCD theorist is less clear, but to see why it will help to consider Cabello's proof.

4. Cabello's proof

Kochen and Specker realized that Gleason's original theorem had combinatoric consequences. Stated in three dimensions, their result says that if each one-dimensional projector represents a unique proposition, then it is impossible to select exactly one member of each orthogonal triple of such projectors. This is because each projector belongs to many orthogonal triples, and so assigning the truth-value "true" to the corresponding proposition has consequences for other triples—consequences that lead to a contradiction in a finite number of steps.

If we restrict ourselves to projectors, this style of argument fails in two dimensions; each projector belongs to exactly one orthogonal pair. If we include POVM elements, the combinatoric situation changes: the number of elements in a POVM can exceed the number of dimensions in the Hilbert space, and a given effect can belong to many POVMs. Cabello considers a family of five 8-element POVMs on two-dimensional Hilbert space and shows that there is no way to select exactly one element from each POVM. However, in a note added in proof, Cabello points out that Masahiro Nakamura found a simpler version of the argument: one that makes use of three 4-element POVMs. We will avail ourselves of this simplification in what follows.

Although Cabello and Nakamura provide geometric interpretations for their constructions, this is not essential. Let A, B and C be any three one-dimensional projectors on H^2, subject only to the constraint that they are distinct and mutually non-orthogonal. Let $A^\perp$, $B^\perp$ and $C^\perp$ denote the projectors orthogonal to A, B and C. Since $1/2X + 1/2X^\perp = 1/2I$ for any projector X, each of the following sets is a POVM:

$$\{1/2A, 1/2A^\perp, 1/2B, 1/2B^\perp\},$$
$$\{1/2A, 1/2A^\perp, 1/2C, 1/2C^\perp\},$$
$$\{1/2B, 1/2B^\perp, 1/2C, 1/2C^\perp\}.$$

There is no set that contains exactly one member from each of these three sets, and so there is no non-contextual way to assign exactly one element from each POVM the value "true". This, in essence, is the proof. Before we assess it, however, we turn to a bookkeeping issue.

5. Ambiguity

In our earlier characterization of NCD HVs, we assumed that each self-adjoint operator represented a single quantity with a unique value. If we attempted to impose this requirement on POVM elements, it would be immediately apparent that the requirement cannot be satisfied. A particularly simple way to see this often comes up in discussion: $\{1/2I, 1/2I\}$ is a POVM. If we require that one element of each POVM be assigned the value "true" and the others the value "false", we are clearly in trouble already. $1/2I$ has no "non-contextual" value even within a single POVM.

This example is not a mere anomaly. Let d and d' be two directions at the same angle θ to the z-axis and consider two sequential measurement experiments. The first consists of a measurement of σ_z followed by a measurement of σ_d. The second consists of a measurement of σ_z followed by $\sigma_{d'}$. Writing P_{z^+} and P_{z^-} for $|z^+\rangle\langle z^+|$ and $|z^-\rangle\langle z^-|$, the POVMs for the two experiments are, respectively,

$$\{|\langle z^+|d^+\rangle|^2 P_{z^+},\ |\langle z^+|d^-\rangle|^2 P_{z^+},\ |\langle z^-|d^+\rangle|^2 P_{z^-},\ |\langle z^-|d^-\rangle|^2 P_{z^-}\}$$

and

$$\{|\langle z^+|d'^+\rangle|^2 P_{z^+},\ |\langle z^+|d'^-\rangle|^2 P_{z^+},\ |\langle z^-|d'^+\rangle|^2 P_{z^-},\ |\langle z^-|d'^-\rangle|^2 P_{z^-}\}.$$

However, both of these are just

$$\{\cos^2(\theta/2)P_{z^+}, \sin^2(\theta/2)P_{z^+}, \sin^2(\theta/2)P_{z^-}, \cos^2(\theta/2)P_{z^-}\}.$$

In spite of this, the two experiments are physically distinct, as is made clear by the fact that they lead to different final states. Even if a POVM contains repeated effects, the POVM does not pick out a single experiment unambiguously. And even if we specify the experiment we have in mind, distinct outcomes may go with identical effects. To stick with the example just sketched, if $\theta = 90°$ then we will have $\cos^2(\theta/2) = \sin^2(\theta/2) = 1/2$, and each POVM will contain repeated elements.

This ambiguity makes clear that effects per se do not represent fundamental quantities. At best, they represent equivalence classes of quantities. However, none of this by itself refutes Busch and Cabello's arguments. It simply reminds us that distinct experimental propositions can be statistically equivalent in all quantum states. The POVM elements encode the statistical information, but they do not distinguish among the statistically equivalent propositions. If we can associate each of Cabello's effects with a univocal experimental proposition, his combinatoric reasoning still stands. And since a measure on effects induces a measure on experimental propositions, Busch's arguments stand as well. Issues about just what experiment a POVM

represents will come up later, however. We turn now to an assessment of Cabello's proof.

5.1. Classical model of Cabello's POVMs

Imagine a box with three drawers. In each drawer is a marble, either black or white; we can assume whatever probabilities we like for the colors of the marbles. Label the drawers A, B and C. And now consider the following three experiments:

Experiment #1. Choose between drawers A and B by flipping a fair coin. If it comes up heads, open drawer A and record the color of the marble. If it comes up tails, open drawer B and record the color of the marble.

Experiment #2. Choose between drawers A and C by flipping a fair coin. If it comes up heads, open drawer A and record the color of the marble. If it comes up tails, open drawer C and record the color of the marble.

Experiment #3. Choose between drawers B and C by flipping a fair coin. If it comes up heads, open drawer B and record the color of the marble. If it comes up tails, open drawer C and record the color of the marble.

Suppose we represent the experiments and their possible outcomes as sets:

$$\{A_W, A_B, B_W, B_B\},\ \{A_W, A_B, C_W, C_B\},\ \{B_W, B_B, C_W, C_B\}.$$

Clearly it is impossible to select exactly one member of each set, but the drawers and their contents are thoroughly classical; there is no meaningful "no go" result here. But now consider the following qubit experiments. One experiment consists in flipping a fair coin and measuring σ_a if the result is heads or σ_b if the result is tails. The other two experiments are similarly defined: for Experiment #2, measure σ_a if the coin comes up heads and σ_c otherwise. For Experiment #3, heads picks out σ_b and tails picks out σ_c. If the quantum state is ρ, the statistics of the experiments are given via the POVMs $\{1/2A, 1/2A^\perp, 1/2B, 1/2B^\perp\}$, $\{1/2A, 1/2A^\perp, 1/2C, 1/2C^\perp\}$, $\{1/2B, 1/2B^\perp, 1/2C, 1/2C^\perp\}$ and the trace rule. But an NCD theorist will have no difficulty with this example. He will say that each spin quantity has a determinate value, with the distribution of those values captured by a measure on an appropriate probability space such as the one in Kochen and Specker's model. The "contextuality" of $1/2A$, $1/2B^\perp$, etc., like the "contextuality" of A_W, B_B, etc., tells us nothing about the underlying quantities. Performing the $\{1/2A, 1/2A^\perp, 1/2B, 1/2B^\perp\}$ experiment and selecting σ_a as the quantity to be measured is compatible with saying that had we chosen the $\{1/2B, 1/2B^\perp, 1/2C, 1/2C^\perp\}$ experiment, the same outcome of the coin toss would have led us to select σ_b. On the NCD view, the fundamental quantities—the determinate spin values—are not contextual at all.

It is still true, of course, that the effects in this example are "contextual". But it would be more accurate to say that the supposed no-go result depends on an equivocation. An effect like "$1/2A$", understood as it has been here, does not pick out an intrinsic feature or fundamental property of the quantum system. In the set $\{1/2A, 1/2A^\perp, 1/2B, 1/2B^\perp\}$, "$1/2A$" means something like "In the context of a choice between measuring σ_a and σ_b, σ_a is chosen and measured with result $+1$". This is a different proposition from "In the context of a choice between measuring σ_a and σ_c, σ_a is chosen and measured with result $+1$", and there is no reason to think that two should have the same truth value, because there is no reason to think the same choice must be made in the two different contexts.

The appeal to coin flips was inspired by various examples in R.W. Spekkens' recent discussion of contextuality [8], but it is used here to make a quite different point. The point is that a mere appeal to the combinatorics of a set of POVMs cannot rule out NCD HVs. But if the elements of POVMs need not be the sorts of things that the NCD theorist would expect to have univocal truth-values, then the case for treating Busch's more sophisticated result as a no-go proof is undercut as well.

The Cabello/Nakamura example is peculiar in another, related way. The argument is formulated in H^2, but consider the following collection of projectors, where each of the factor spaces is two-dimensional:

$$A \otimes I \otimes I,\ A^\perp \otimes I \otimes I,\ I \otimes B \otimes I,\ I \otimes B^\perp \otimes I,$$
$$I \otimes I \otimes C,\ I \otimes I \otimes C^\perp.$$

Here A and $A^\perp$ are orthogonal one-dimensional projectors on the first Hilbert space, and similarly for B, $B^\perp$, etc. The six projectors above are mutually commuting. For simplicity, drop the tensor products and denote $A \otimes I \otimes I$ by A_1, and so on. Clearly, $(A_1 + A_1^\perp)$ is the identity on the eight-dimensional space, as are $(B_2 + B_2^\perp)$ and $(C_3 + C_3^\perp)$. Therefore, each of the following sets is a POVM:

$$\{1/2A_1, 1/2A_1^\perp, 1/2B_2, 1/2B_2^\perp\},$$
$$\{1/2A_1, 1/2A_1^\perp, 1/2C_3, 1/2C_3^\perp\},$$
$$\{1/2B_2, 1/2B_2^\perp, 1/2C_3, 1/2C_3^\perp\}.$$

Cabello's proof goes through for this collection of POVMs: it is impossible to select exactly one member from each set. But quantum theory itself makes it implausible that this entails anything deep about the quantities involved. After all, if the state of the joint system is any of $|a+\rangle \otimes |b+\rangle \otimes |c+\rangle$, etc., the results of any of the three experiments are completely accounted for by the quantum state and the outcome of the coin toss. Quantum theory, it seems, provides its own "hidden variable" account of this case.

6. Further questions

As we have seen, the NCD theorist has no reason to be embarrassed by the Cabello/Nakamura example. Of course, not every POVM can be thought of as merely presenting us with a weighted choice between ordinary projective measurements. But now we run into a different problem: deciding when it is reasonable to require the NCD theorist to assign definite values to experimental propositions in the first place.

Consider sequential measurements once again. Here we get more than two results not because of an externality like a coin toss that chooses between two experiments, but because we have linked together measurements of quantities that the NCD theorist would presumably regard as intrinsic to the qubit: spin in different directions. Sequential measurements can be represented by POVMs, but would a hidden variable theorist have to agree that the corresponding experimental propositions are assigned well-defined truth values by the hidden variables?

It is not obvious that he would. As we pointed out earlier, any plausible NCD theory will claim that measurement changes the state of a system, and there is no reason why a non-contextual *determinate* theory should be compelled to treat that change as deterministic. The NCD theorist can say that before the z-spin measurement, the hidden variable assigns a determinate value to the z-spin and fixes the outcome of the corresponding measurement. Thus, if the qubit enters the z^+ channel, this is because it had spin up in direction z. However, the z-spin measurement will change the hidden state so that the new distribution of spin values fits the quantum statistics of the state $|z^+\rangle$. The resulting d-spin value will determine which channel the qubit enters in the final stage of the experiment, and the NCD theorist has no reason to concede that the change that gave rise to the new d-spin value was deterministic. If it was not, there is no fact in advance of the actual experiment about whether an individual qubit will emerge from the d^+ channel or the d-channel, and no reason why the NCD theory should posit one. Indeed, as we saw above, Kochen and Specker's model of the spin-$1/2$ system incorporates precisely this sort of indeterminism and can take this case in stride.

7. Conclusion

The considerations offered here do not simply let the NCD theorist off the hook. For the Cabello–Nakamura example and the case of sequential measurements, it was easy to see how the NCD theorist could construct a positive account of the experiments—either by understanding the POVM in terms of a choice among projective measurements or by invoking, as Kochen and Specker's model does, an indeterministic account of state change from one stage of an experiment to another. This does not exhaust the range of generalized experiments on the qubit. It may well be that by considering more elaborate examples—perhaps especially ones that depend in an essential way on interference effects—we could come up with cases that do not permit a reasonable NCD model. The thesis of this Letter is that purely formal results such as Busch's and Cabello's cannot settle the question. This leaves many open questions about the location of the boundary between classical and quantum and about the foundational role of POVMs, but those are problems for another occasion.

Acknowledgements

I would like to thank Jeffrey Bub for helpful discussion and Paul Busch for useful comments on an earlier draft. I would also like to thank the General Research Board of the University of Maryland for support during the preparation of this Letter.

References

[1] S.J. van Enk, Phys. Rev. Lett. 84 (2000) 789.
[2] P. Busch, Phys. Rev. Lett. 91 (2003) 120403.
[3] A.M. Gleason, J. Math. Mech. 6 (1957) 885.
[4] A. Cabello, Phys. Rev. Lett. 90 (2003) 190401.
[5] S. Kochen, E.P. Specker, J. Math. Mech. 17 (1967) 59.
[6] J. von Neumann, Mathematische Grundlagen der Quantenmechanik, Springer, Berlin, 1932.
[7] J.S. Bell, Rev. Mod. Phys. 38 (1966) 447.
[8] R.W. Spekkens, Phys. Rev. A 71 (2005) 052108.