Hidden Variables and Commutativity in Quantum Mechanics

Benjamin Feintzeig

Department of Logic and Philosophy of Science, University of California, Irvine

bfeintze@uci.edu

January 1, 2013

Abstract

This paper takes up a suggestion that the reason we cannot find hidden variable theories for quantum mechanics, as in Bell's Theorem, is that we require them to assign joint probability distributions on incompatible observables; these joint distributions are empirically meaningless on one standard interpretation of quantum mechanics. Some have proposed to get around this problem by using *generalized probability spaces*. I present a "no-go" theorem to show a sense in which generalized probability spaces can't serve as hidden variable theories for quantum mechanics, so the proposal for getting around Bell's Theorem fails.

Contents

1 Introduction

Various well known "no-go" theorems purport to show that one cannot find hidden variable theories (HVTs) for all quantum mechanical experiments. Some have suggested that the reason we cannot find HVTs is that we expect too much of them. On one standard way of understanding HVTs, we require assignments of joint probabilities to all pairs of observables, but the usual interpretation of quantum mechanics tells us that certain observables, namely non-commuting ones, cannot be measured at the same time. Joint probability distributions on non-commuting observables lack empirical meaning, which motivates an investigation of alternative ways of constructing HVTs.

This paper explores an alternative way of constructing HVTs[1] that does just this—while HVTs are sometimes understood as *classical probability spaces*[2], one can weaken the usual axioms in a particular way to obtain a definition of *generalized probability spaces*, in which we allow ourselves to forgo joint probabillity assignments on non-commuting observables. In section 2, we review the "no-go" theorems concerning classical probability spaces. In section 3, we motivate the consideration of generalized probability spaces by showing that the problematic cases, i.e. the quantum mechanical experiments that cannot be given a classical probability space representation, all have non-commuting observables. In section 4, we explore some strange properties of generalized probability spaces as candidates for HVTs. In section 5, we show that one cannot find generalized probability space representations for all quantum mechanical experiments unless we give up very natural constraints on our HVTs. We conclude with a discussion of how the main result of this paper reveals an unexpected connection between Bell's theorem and the Kochen-Specker theorem.

[1]The HVTs considered in this paper are all so-called *non-contextual* HVTs in that they assign definite values to all possible properties of the experimental system at the same time. We don't consider Bohmian mechanics, a HVT that only assigns definite values to observables that are a function of position.

[2]I show the relationship between classical probability spaces and some more well-known ways of thinking about HVTs in section 2.

2 Bell's Theorem and Classical Probability Spaces

Definition 1. For the purposes of this paper, we take a *quantum mechanical experiment* to be an ordered triple $(\mathcal{H}, \psi, \mathfrak{S})$, where $\mathcal{H}$ is a Hilbert space, $\psi \in \mathcal{H}$ is a unit vector (i.e. $\langle \psi, \psi \rangle = 1$), and $\mathfrak{S} = \langle P_1, ..., P_n \rangle$ is an ordered sequence of projection operators[3] onto subspaces of $\mathcal{H}$.

The vector ψ represents the quantum state in which we have prepared the experimental system, and each projection operator P_i corresponds to a "yes-no" measurement, i.e. a measurement with exactly two possible outcomes. We assign a probability value, p_i, to the event of obtaining a "yes" outcome for the measurement corresponding to P_i as follows:

$$p_i = \langle \psi, P_i \psi \rangle$$

Furthermore, if two projection operators P_i, P_j are compatible (i.e. they commute: $[P_i, P_j] = 0$), then we can measure them together, and so quantum mechanics ascribes a joint probability, p_{ij}, to the event of obtaining a "yes" outcome for both measurements:

$$p_{ij} = \langle \psi, P_i P_j \psi \rangle$$

A quantum mechanical experiment brings with it a data set of type (n, S), where n is the number of operators in $\mathfrak{S}$, and $S = \{\langle i, j \rangle : [P_i, P_j] = 0 \text{ and } 1 \leq i \leq j \leq n\}$. We yield a set of predictions, or a *probability data set*, the $(n + |S|)$-tuple $\langle p_1, ..., p_n, ..., p_{ij}, ... \rangle$, where p_{ij} appears if and only if $\langle i, j \rangle \in S$, i.e. if P_i and P_j are compatible, and the p_{ij} terms are ordered lexicographically.

First, we give a standard presentation of Bell's "no-go" theorem for HVTs before moving on to Pitowsky's slightly more abstract variation in terms of classical probability spaces.

[3]For the purposes of this paper, we'll supose $\mathfrak{S}$ is finite, although in general it need not be.

2.1 Bell's Derivation of the Bell inequalities

The usual understanding of Bell's theorem takes the settings of our measurement apparatus into account in addition to the measurement outcomes[4]. We can write the above probabilistic predictions as conditional on measurement settings: let $pr_{QM}(A_1, ..., A_n|a_1, ..., a_n)$ represent the probability that we obtain outcome A_i on the ith measurement apparatus given that it was prepared with measurement setting a_i. For example, if we consider the EPR setup[5], in which two photons are emitted in opposite directions in the singlet state from a common source and a polarizer sheet is placed at each of the right and left ends of the setup, then $n = 2$, each a_i represents the direction of polarization of the polarizer sheet which can be rotated in a plane, and each A_i can take the values of either "yes"—the photon passed through the polarizer sheet—or "no"—the photon did not pass through the polarizer sheet. We determine these probabilities from the Hilbert space associated with the experimental system using the above prescription.

Definition 2. A *Bell HVT* for a quantum mechanical experiment with n measurement apparatuses is an ordered triple $(X, \{pr_{HV}(A_1, ..., A_n|a_1, ..., a_n; x) : x \in X\}, \rho)$, where X is a set of states, each state $x \in X$ determines a probability function pr_{HV} of the form displayed, and ρ is a probability density (i.e. $\rho : X \to [0, 1]$ and $\int_X \rho(x)dx = 1$) such that[6]

$$pr_{QM}(A_1, ..., A_n|a_1, ..., a_n) = \int_X pr_{HV}(A_1, ..., A_n|a_1, ..., a_n; x)\rho(x)dx$$

One understands X to be the set of hidden states, and the probability functions give us new predictions based on the further information or hidden variables associated with one of those hidden states. The difference between quantum mechanical predictions and predictions of a hidden variable theory is that probabilities determined by the hidden variable theory are conditional not only on the measurement settings, but also on the hidden state x. The probability density $\rho(x)$ represents the probability of finding the experimental system in the hidden state x.

[4]This section follows the presentation in Malament (2012, section 1).

[5]See Einstein, Podolsky and Rosen (1935), and Bohm and Aharanov (1957).

[6]Although we have not specified the operation of integration, for what follows we need only assume that the integral has certain standard properties that all kinds of integrals possess, such as additivity.

Now we put constraints on the the probability functions of the HVT—for our purposes only two constraints are relevant. We formulate the constraints in terms of the EPR setup in which $n = 2$.

Quasi-determinateness: For all A_1, A_2, a_1, a_2, x, $pr_{HV}(A_1, A_2|a_1, a_2; x) = 0$ or 1.

Locality: For all $A_1, A_2, a_1, a_1', a_2, a_2', x$,

$$pr_{HV}(A_1, \_|a_1, a_2; x) = pr_{HV}(A_1, \_|a_1, a_2'; x)$$
$$pr_{HV}(\_, A_2|a_1, a_2; x) = pr_{HV}(\_, A_2|a_1', a_2; x)$$

Quasi-determinateness requires the existence of a definite value for each property: "yes" if the probability is 1 or "no" if the probability is zero. Locality requires that the measurement outcomes *here* cannot depend on the settings *there*, where *here* and *there* are two distinct (possibly spacelike separated) apparatuses; the outcomes *here* only depend on the settings *here*.

From the constraints of Quasi-determinateness and Locality, one can derive a characteristic inequality, known as Bell's inequality[7]. So if all quantum mechanical experiments had Local, Quasi-determinate HVTs, then the probabilistic predictions of each of those quantum mechanical experiments would satisfy Bell's inequality. But when we consider the quantum mechanical experiment that represents the EPR setup, in which two photons or electrons are emitted in the singlet state and we take the appropriate measurements of polarization or spin, we find that the probability data set corresponding to that experimental system violates Bell's inequality (Bell 1964, p. 198; Pitowsky 1989, p. 84). So we conclude that there could not be a Local, Quasi-determinate hidden variable theory for all quantum mechanical experiments—in particular, not for the EPR setup.

Theorem 1. (Bell) There are quantum mechanical experiments for which there does not exist a Local, Quasi-determinate Bell HVT.

[7] In our notation, Bell's inequality takes the form

$$0 \leq pr_{QM}(yes, \_|a, \_) + pr_{QM}(\_, yes|\_, b) + pr_{QM}(yes, yes|a', b') - pr_{QM}(yes, yes|a, b') - pr_{QM}(yes, yes|a', b) - pr_{QM}(yes, yes|a, b) \leq 1$$

I've deemphasized the importance of this inequality because it is not relevant to the "no-go" theorems we'll see later.

2.2 Pitowsky's Derivation of the Bell Inequalities

Pitowsky's presentation of the Bell inequalities begins by thinking of HVTs in terms of classical probability spaces:

Definition 3. A *σ-algebra* Σ on a set X is a non-empty set of subsets of X such that for all $A, B \subseteq X$

(i) If $A \in \Sigma$, then $(X - A) \in \Sigma$, and

(ii)[8] If $A, B \in \Sigma$, then $A \cup B \in \Sigma$.

Definition 4. A *classical probability space* (Gudder 1988, p. 2; Krantz et. al, p. 200; Billingsley 1979, p. 19) is an ordered triple (X, Σ, μ), where X is a non-empty set of states, Σ is a σ-algebra of subsets of X, and $\mu : \Sigma \to \mathbb{R}$ is a real valued function such that:

(i) $\mu(X) = 1$,

(ii) $\mu(A) \geq 0$, and

(iii) If $A \cap B = \varnothing$, then $\mu(A \cup B) = \mu(A) + \mu(B)$.

We still think of the elements of X as the complete, or hidden, states of the experimental system, and each measurable set in Σ corresponds to a property that the experimental system can either have or not have by falling in or out of that set. If we knew the hidden state of our experimental system, we would know all of its properties, but since we don't, we can only ascribe probability values to the properties of our experimental system by assigning a measure to subsets of our space.

There are two important facts that will help us distinguish classical probability spaces from the generalized probability spaces we consider later on. First, in a classical probability space, if $A, B \in \Sigma$, then $(X - ((X - A) \cup (X - B))) = A \cap B \in \Sigma$. Thus we are required to assign a probability value to the intersection of any two sets we assign probability values to individually. Whenever we assign probabilities to two individual properties, such as being in A and being in B, we assign a

[8] We restrict our attention in this paper to finite unions although in general, a σ-algebra allows for countable unions, and the axioms of the probability space change accordingly. Technically, what I call σ-algebras here are mere algebras, and what I call σ-additive classes (in section 4) are mere additive classes.

probability to the conjunction of those properties, being in A and B at the same time.

Second, it is easy to show that in a classical probability space, for every $A, B \in \Sigma$, $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$. One can generalize this formula by induction to obtain the *inclusion-exclusion formula* (Billingsley 1979, p. 20):

$$\mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i) - \sum_{i<j} \mu(A_i \cap A_j) + \sum_{i<j<k} \mu(A_i \cap A_j \cap A_k) + ... + (-1)^{n+1} \mu(A_1 \cap ... \cap A_n)$$

From this it follows that the union of a finite number of probability zero sets cannot have anything but probability zero. In particular, the union of a finite number of probability zero sets cannot equal the whole space X, because then the union would have probability one. We'll see how generalized probability spaces differ in these properties in section 4.

There are two ways in which we can connect quantum mechanical experiments to classical probability spaces, as in the following definitions.

Definition 5. A quantum mechanical experiment $(\mathcal{H}, \psi, \mathfrak{S})$ has a *restricted classical probability space representation* iff there is a classical probability space (X, Σ, μ) with sets $A_1, ..., A_n \in \Sigma$ such that for all P_i, P_j,

$$\mu(A_i) = p_i = \langle \psi, P_i \psi \rangle$$

and if $[P_i, P_j] = 0$, then

$$\mu(A_i \cap A_j) = p_{ij} = \langle \psi, P_i P_j \psi \rangle$$

In Definition 5, we only considered the case where two observables commute, and imposed a natural requirement amounting to a statement that we represent their joint probabilities in the usual way. In general, we may have more than two observables that all pairwise commute, and in this case they are all measurable simultaneously. This idea is captured in the following definition.

Definition 6. A quantum mechanical experiment $(\mathcal{H}, \psi, \mathfrak{S})$ has a *full-blown classical probability space representation* iff there is a classical probability space (X, Σ, μ) with sets $A_1, ..., A_n \in \Sigma$ such that for all P_i,

$$\mu(A_i) = p_i = \langle \psi, P_i \psi \rangle$$

and for all $P_i, P_j, ..., P_k$ that are compatible (i.e. they pairwise commute), the corresponding sets $A_i, A_j, ..., A_k$ satisfy

$$\mu(A_i \cap A_j \cap ... \cap A_k) = p_{ij...k} = \langle \psi, P_i P_j ... P_k \psi \rangle$$

Why should we care about classical probability space representations? One might argue they are irrelevant to Bell's theorem because there is no mention of Locality in their construction. But every Local, Quasi-determinate Bell HVT leads to a classical probability space representation in the following way (Malament 2012, p. 7).

Proposition 1. If $(X, \{pr_{HV}(A_1, A_2|a_1, a_2; x) : x \in X\}, \rho)$ constitutes a Local, Quasi-determinate Bell HVT for a specified quantum mechanical experiment, then there is a classical probability space (X, Σ, μ) such that for all a_1, a_2, there are sets[9] $L_{a_1}, R_{a_2}, \in \Sigma$ for which[10]

$$\mu(L_{a_1}) = pr_{QM}(yes, \_|a_1, \_)$$

$$\mu(R_{a_2}) = pr_{QM}(\_, yes|\_, a_2)$$

$$\mu(L_{a_1} \cap R_{a_2}) = pr_{QM}(yes, yes|a_1, a_2)$$

[9] L_{a_1} is the set of hidden states which will yield a "yes" outcome on the left, given setting a_1, and R_{a_2} is the set of hidden states which will yield a "yes" outcome on the right, given setting a_2.

[10] I.e. the specified quantum mechanical experiment has a (both restricted and full-blown) classical probability space representation.

Proof. For any a_1, a_2, let

$$L_{a_1} = \{x \in X : pr_{HV}(yes, \_|a_1, \_) = 1\}$$
$$R_{a_2} = \{x \in X : pr_{HV}(\_, yes|\_, a_2) = 1\}$$

Notice that the expressions on the right hand side are only well-defined because we've assumed Locality. From this, it follows that

$$L_{a_1} \cap R_{a_2} = \{x \in X : pr_{HV}(yes, yes|a_1, a_2) = 1\}$$

Let Σ be the σ-algebra generated by all of the sets of the form L_{a_1} and R_{a_2}. Define μ for each measurable set $C \in \Sigma$ by

$$\mu(C) = \int_C \rho(x)dx$$

It follows that

(i) $\mu(X) = \int_X \rho(x)dx = 1$

(ii) $\mu(C) = \int_C \rho(x)dx \geq 0$, and

(iii) If $C_1 \cap C_2 = \varnothing$, then $\mu(C_1 \cup C_2) = \int_{C_1 \cup C_2} \rho(x)dx = \int_{C_1} \rho(x)dx + \int_{C_2} \rho(x)dx = \mu(C_1) + \mu(C_2)$.

Therefore (X, Σ, μ) is a classical probability space. Furthermore,

$$pr_{QM}(yes, \_|a_1, \_) = \int_X pr_{HV}(yes, \_|a_1, \_; x)\rho(x)dx = \int_{L_{a_1}} 1 \cdot \rho(x)dx + \int_{X - L_{a_1}} 0 \cdot \rho(x)dx = \mu(L_{a_1})$$

$$pr_{QM}(\_, yes|\_, a_2) = \int_X pr_{HV}(\_, yes|\_, a_2; x)\rho(x)dx = \int_{R_{a_2}} 1 \cdot \rho(x)dx + \int_{X - R_{a_2}} 0 \cdot \rho(x)dx = \mu(R_{a_2})$$

$$\begin{aligned} pr_{QM}(yes, yes|a_1, a_2) &= \int_X pr_{HV}(yes, yes|a_1, a_2; x)\rho(x)dx \\ &= \int_{L_{a_1} \cap R_{a_2}} 1 \cdot \rho(x)dx + \int_{X-(L_{a_1} \cap R_{a_2})} 0 \cdot \rho(x)dx \\ &= \mu(L_{a_1} \cap R_{a_2}) \end{aligned}$$

□

Since every Local, Quasi-determinate Bell HVT comes with a classical probability space representation, showing that there are quantum mechanical experiments with no classical probability space representation would imply that there are quantum mechanical experiments with no Local, Quasi-determinate Bell HVT. We now turn our attention to whether there are classical probability space representations for all quantum mechanical experiments.

Given a quantum mechanical experiment with a probability data set of type (n, S), following the formalism of Pitowsky (1989, p. 21), we let $I = \{0, 1\}^n$ and $\epsilon = \langle \epsilon_1, ..., \epsilon_n \rangle \in I$ be an n-tuple of zeros and ones. Let $p^\epsilon = \langle \epsilon_1, ..., \epsilon_n, ..., \epsilon_i\epsilon_j, ... \rangle \in \mathbb{R}^{n+|S|}$, where the product $\epsilon_i\epsilon_j$ appears just in case $\langle i, j \rangle \in S$. Let $c(n, S)$ be the closed, convex polytope in $\mathbb{R}^{n+|S|}$ whose vertices are the 2^n vectors of the form p^ϵ, i.e. $c(n, S)$ contains all vectors of the form $u = \sum_{\epsilon \in I} \lambda(\epsilon)p^\epsilon$, where each $\lambda(\epsilon)$ is a non-negative scalar and $\sum_{\epsilon \in I} \lambda(\epsilon) = 1$.

Theorem 2. (Pitowsky 1989, p. 22) A probability data set $\langle p_1, ..., p_n, ..., p_{ij}, ... \rangle$ of type (n, S) has a restricted[11] classical probability space representation iff $\langle p_1, ..., p_n, ..., p_{ij}, ... \rangle \in c(n, S)$.

When a probability data set belongs to this characteristic polytope, it satisfies a certain set of inequalities describing the bounding surfaces of $c(n, S)$—one of these is Bell's inequality. Once again, it is well-known that the probability data sets for some quantum mechanical experiments, namely the EPR setup, violate these inequalities. So we conclude with a "no-go" theorem for classical probability space representations as HVTs.

[11] It seems that one could generalize this result to cover full-blown classical probability space representations, but this is beyond the scope of this paper. See footnote 12 for more.

Theorem 3. (Pitowsky) There are quantum mechanical experiments for which there are no restricted (and hence no full-blown) classical probability space representations.

Pitowsky's theorem implies Bell's theorem via Proposition 1. If a witness to Pitowsky's theorem had a Local, Quasi-determinate Bell HVT, then it would also have a classical probability space representation, so by modus tollens it does not have the former kind of HVT. A natural question to ask is whether the implication holds in the opposite direction—does Bell's theorem imply Pitowsky's theorem? It seems that classical probability space representations lack the necessary information to construct a Bell HVT because they encode only information about measurement outcomes and not measurement settings. We can think of Bell HVTs as a specialization of classical probability space representations, where we add extra information about distinct measurement apparatuses and their settings, which allows us to formulate the Locality constraint[12].

3 Incompatible Observables

Arthur Fine argues that the reason we cannot find HVTs in the form of classical probability spaces for quantum mechanical experiments is that we are trying to assign probabilities to conjunctions of measurements on incompatible observables, and these probability assignments have no empirical meaning since we cannot measure incompatible observables simultaneously. Fine writes,

> "...hidden variables and the Bell inequalities are all about...imposing requirements to make well defined precisely those probability distributions for noncommuting observables whose rejection is the very essence of quantum mechanics" (Fine 1982a, p. 294).

By representing our experimental system in a classical probability space, we force ourselves to assign probability values to the conjunctions of all outcomes that we assign probabilities to individually, even if those outcomes correspond to incompatible observables. One might have noticed Definitions 5 and 6 only put constraints on the measures we assign to intersections of sets corresponding to

[12] Pitowsky (1989, p. 92) gives a nice description of how we might understand violations of locality as manifestations of violating the merely logical constraints of a classical probability space representation.

compatible observables. But if P_i and P_j are incompatible observables and we represent them in a classical probability space by sets $A_i, A_j \in \Sigma$, then it follows that $A_i \cap A_j \in \Sigma$ so we require ourselves to assign *some probability or other* to the outcomes of measuring incompatible observables simultaneously. But quantum mechanics tells us that incompatible observables cannot be measured at the same time, so at the very least it seems strange to assign a probability to the joint outcome.

In two papers (Fine 1982a; Fine 1982b), Fine presents a number of technical results[13] concerning joint distributions and compatible observables to motivate these claims. Here, we present yet another variant, using Pitowsky's powerful results.

Theorem 4. For all quantum mechanical experiments $(\mathcal{H}, \psi, \mathfrak{S})$, if all of the projection operators $P_1, ..., P_n$ are compatible (i.e. for all $i, j \leq n$, $[P_i, P_j] = 0$), then the experimental system has a restricted[14] classical probability space representation.

Proof.[15] Suppose all of the projection operators P_i, P_j commute. We show that the vector $\langle p_1, ..., p_n, ..., p_{ij}, ...\rangle$ defined by the above values belongs to $c(n, S)$.

For any P_i, let $P_i^1 = P_i$ and $P_i^0 = \mathbb{I} - P_i$. Let $P(\epsilon) = P_1^{\epsilon_1} \cdot ... \cdot P_i^{\epsilon_i} \cdot ... \cdot P_n^{\epsilon_n}$.

Notice each $P(\epsilon)$ is a projection operator since we have assumed all P_i, P_j commute, and clearly, $\sum_{\epsilon \in I} P(\epsilon) = \mathbb{I}$.

Furthermore, if $\epsilon \neq \epsilon'$, then $P(\epsilon)$ and $P(\epsilon')$ differ for some P_i and since the P_i's are commutative, it follows that $P(\epsilon)P(\epsilon') = ... \cdot P_i^0 \cdot P_i^1 \cdot ... = ... \cdot P_i \cdot (\mathbb{I} - P_i) \cdot ... = 0$.

Let $\lambda(\epsilon) = \langle \psi, P(\epsilon)\psi \rangle$. For all $\epsilon \in I$, $\lambda(\epsilon) \geq 0$ since $P(\epsilon)$ is Hermitian. Furthermore, $\sum_{\epsilon \in I} \lambda(\epsilon) = \sum_{\epsilon \in I} \langle \psi, P(\epsilon)\psi \rangle = \langle \psi, \sum_{\epsilon \in I} P(\epsilon)\psi \rangle = \langle \psi, \psi \rangle = 1$.

For all i,$j \leq n$, $P_i = \sum_{\{\epsilon \in I : \epsilon_i = 1\}} P(\epsilon)$ and $P_i P_j = \sum_{\{\epsilon \in I : \epsilon_i = \epsilon_j = 1\}} P(\epsilon)$.

Hence, $p_i = \langle \psi, P_i \psi \rangle = \sum_{\{\epsilon \in I : \epsilon_i = 1\}} \langle \psi, P(\epsilon)\psi \rangle = \sum_{\{\epsilon \in I : \epsilon_i = 1\}} \lambda(\epsilon) = \sum_{\epsilon \in I} \lambda(\epsilon)\epsilon_i = \sum_{\epsilon \in I} \lambda(\epsilon)(p^\epsilon)_i$, and $p_{ij} = \langle \psi, P_i P_j \psi \rangle = \sum_{\{\epsilon \in I : \epsilon_i = \epsilon_j = 1\}} \langle \psi, P(\epsilon)\psi \rangle = \sum_{\{\epsilon \in I : \epsilon_i = \epsilon_j = 1\}} \lambda(\epsilon) = \sum_{\epsilon \in I} \lambda(\epsilon)\epsilon_i\epsilon_j = \sum_{\epsilon \in I} \lambda(\epsilon)(p^\epsilon)_{ij}$.

[13]In particular see Theorem 7 in Fine 1982b, p. 1309.

[14]If one generalized Pitowsky's result in Theorem 2, as we conjectured in footnote 9 one might be able to do, then we conjecture further that one would be able to generalize Theorem 4 for full-blown classical probability space representations.

[15]This is analogous to the proof of the "left-to-right" direction of Theorem 1 found in Pitowsky (1989, p. 23).

Thus, $\langle p_1, ..., p_n, ..., p_{ij}, ...\rangle \in c(n, S)$, so by Theorem 2, the desired classical probability space representation exists. $\square$

Since we have no trouble finding HVTs, in the form of classical probability spaces when our experimental system involves only compatible observables, one might claim that the reason we fail to find classical probability space representations for quantum mechanical experiments is that we insist on assigning a probability to the conjunction of measurements of incompatible observables when such assignments are meaningless. The next sections deal with an attempt to construct a formalism in which we allow ourselves to forgo these problematic joint probability assignments.

4 Generalized Probability Spaces

Definition 7. A *σ-additive class* (Gudder 1988, 90) Σ on a set X is a non-empty set of subsets of X such that for all $A, B \subseteq X$

(i) If $A \in \Sigma$, then $(X - A) \in \Sigma$, and

(ii) If $A, B \in \Sigma$ and $A \cap B = \varnothing$, then $A \cup B \in \Sigma$.

Definition 8. A *generalized probability space* (Krantz et al. 1971, p. 214; Gudder 1988, p. 169) is an ordered triple (X, Σ, μ), where X is a non-empty set of states, Σ is a σ-additive class of subsets of X, and $\mu : \Sigma \to \mathbb{R}$ is a real valued function such that:

(i) $\mu(X) = 1$,

(ii) $\mu(A) \geq 0$, and

(iii) If $A \cap B = \varnothing$, then $\mu(A \cup B) = \mu(A) + \mu(B)$.

How are generalized probability spaces different from classical probability spaces? The only difference comes in part (ii) of Definition 7. In a generalized probability space we may have $A, B \in \Sigma$, but if A and B are not disjoint, then it is possible that $A \cup B \notin \Sigma$ and $A \cap B \notin \Sigma$. While in a classical probability space we were required to assign probability values to the conjunction of any two outcomes we assigned probabilities to individually, even if they were incompatible, in a generalized probability space we do not require this. This makes generalized probability spaces a prime candidate for study with respect to quantum mechanical experiments, since we can refrain from assigning a probability to the intersections of sets corresponding to incompatible observables.

There is a natural reading of Fine's work in which he is interpreted as advocating the use of generalized probability spaces for quantum mechanics. This is suggested by the following remarks:

> "Perhaps, then, we ought to accept the straight-line induction; that where ... quantum mechanics does not give a well-defined joint distribution, neither would experiments. After all, if we hold that probabilities (including joint probabilities) are real properties, then some observables may simply not have them" (Fine 1982b, p. 1310).

Fine, however, never explicitly addresses generalized probability spaces as they've been defined here. On the other hand, many others have made their support for the use of generalized probability spaces in quantum mechanics explicit—here are some examples:

Krantz et al. write,

> "The notions of event and probability given in [the definition of a classical probability space] have proved satisfactory for almost all scientific purposes. The one outstanding exception is quantum mechanics. In that theory both [$\mu(A)$] and [$\mu(B)$] may exist and yet [$\mu(A \cap B)$] need not" (Krantz et al. 1971, p. 214).

Suppes writes,

> "[T]here is no joint probability distribution of position and momentum... there is no possibility of measuring them jointly at all, because their joint distribution does not exist" (Suppes 1963, p.335).

And he elaborates in a later paper,

> "[T]he joint probability of two events does not necessarily exist in quantum mechanics... Roughly speaking,...the probability distribution of a single quantum-mechanical random variable is classical, and the deviations arise only when several random variables or different kinds of events are considered... [Generalized][16] probability spaces can be used as the basis for an axiomatic development of classical quantum mechanics[17]" (Suppes 1966, p. 345-347).

And Gudder writes,

> "[M]uch of quantum mechanics can be described in the framework of σ-additive classes... This will ultimately result in a framework for a general theory of quantum probability spaces" (Gudder 1988, 169).

The rest of this paper explores the properties of generalized probability spaces, and the extent to which we can or cannot use them to represent quantum mechanical experiments.

In generalized probability spaces, since unions and intersections of measurable sets are not required to be measurable, the inclusion-exclusion formula does not hold in general. However, we have some slightly weaker results.

Lemma 1: If (X, Σ, μ) is a generalized probability space, A,$B \in \Sigma$, and $C = A \cap B$, then the following are equivalent:

(1) $A \cup B \in \Sigma$

(2) $(A - C) \in \Sigma$ and $(B - C) \in \Sigma$

(3) $(A - C) \in \Sigma$

(4) $C \in \Sigma$.

[16]Suppes refers to the objects that I've defined to be generalized probability spaces as "quantum-mechanical probability spaces."

[17]Later in his paper, Suppes makes a further abstraction to non-Boolean versions of generalized probability spaces, but here I'll deal only with the Boolean structures defined above.

Proof: (1⇒2) Suppose $A \cup B \in \Sigma$. Then $X - A \cup B \in \Sigma$ is disjoint from B so $(X - A \cup B) \cup B \in \Sigma$ and hence $(A - C) = X - ((X - A \cup B) \cup B) \in \Sigma$. Similarly, $(B - C) \in \Sigma$.

(2⇒3) Trivial.

(3⇒4) Suppose $(A - C) \in \Sigma$. Then $X - A$ is disjoint from $A - C$ so $(X - A) \cup (A - C) \in \Sigma$, and hence $X - ((X - A) \cup (A - C)) = A \cap B \in \Sigma$.

(4⇒1) Suppose $C = A \cap B \in \Sigma$. Then since $X - A$ is disjoint from $A \cap B$, $(X - A) \cup (A \cap B) \in \Sigma$, so $A - C = X - ((X - A) \cup (A \cap B)) \in \Sigma$. Similarly, $(B - C) \in \Sigma$. Since $A - C$, $B - C$, and C are all disjoint, $A \cup B = (A - C) \cup C \cup (B - C) \in \Sigma$. □

Lemma 2: If (X, Σ, μ) is a generalized probability space and A,$B \in \Sigma$, then if either $A \cap B$ or $A \cup B$ is in Σ, then

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$$

Proof. If either $A \cap B$ or $A \cup B$ are in Σ, then by Lemma 1, they both are. The proof then proceeds exactly as for a classical probability space. □

Notice that if we have two measurable sets A and B such that $A \cap B \in \Sigma$ and $\mu(A \cap B) = 0$, then $\mu(A \cup B) = \mu(A) + \mu(B)$.

Lemma 3: If (X, Σ, μ) is a generalized probability space and $A_1, ..., A_n \in \Sigma$, then if for all $B \subseteq \{A_1, ..., A_n\}$, $\bigcap B \in \Sigma$, then $\bigcup_{i=1}^{n} A_i \in \Sigma$ and

$$\mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i) - \sum_{i<j} \mu(A_i \cap A_j) + \sum_{i<j<k} \mu(A_i \cap A_j \cap A_k) + ... + (-1)^{n+1} \mu(A_1 \cap ... \cap A_n) \quad (1)$$

Proof. By induction. Base step: If $n = 2$, then we just have a restatement of Lemma 2.
Induction step: Suppose the claim holds for n. Then consider $A_1, ..., A_{n+1} \in \Sigma$. Suppose that for all $B \subseteq \{A_1, ..., A_{n+1}\}$, $\bigcap B \in \Sigma$.

First notice that $\bigcup_{i=1}^{n} A_i \in \Sigma$ satisfies formula (1) by the induction hypothesis.

Next, we claim $((\bigcup_{i=1}^{n} A_i) \cap A_{n+1}) = \bigcup_{i=1}^{n}(A_i \cap A_{n+1}) \in \Sigma$. To see this, let $C_i = A_i \cap A_{n+1}$ and $B' \subseteq \{C_1, ..., C_n\}$. Then it follows that $\bigcap B' \in \Sigma$ because it is just the intersection of some A_i's (for $i < n$) intersected with A_{n+1}. In other words $\bigcap B' = A_i \cap A_j \cap ... \cap A_{n+1}$ (for some $i, j, ...$) is one of our original $\bigcap B$'s, so it is measurable. It follows by the induction hypothesis that $\bigcup_{i=1}^{n} C_i = ((\bigcup_{i=1}^{n} A_i) \cap A_{n+1}) \in \Sigma$ satisfies formula (1).

By Lemma 2, since $((\bigcup_{i=1}^{n} A_i) \cap A_{n+1}) \in \Sigma$, $((\bigcup_{i=1}^{n} A_i) \cup A_{n+1}) = \bigcup_{i=1}^{n+1} A_i \in \Sigma$ satisfies the following formula:

$$\mu(\bigcup_{i=1}^{n+1} A_i) = \mu(\bigcup_{i=1}^{n} A_i) + \mu(A_{n+1}) - \mu(\bigcup_{i=1}^{n}(A_i \cap A_{n+1}))$$

Plugging in the expressions for $\mu(\bigcup_{i=1}^{n} A_i)$ and $\mu(\bigcup_{i=1}^{n}(A_i \cap A_{n+1}))$ from formula (1) yields the desired result for $n + 1$. $\square$

Notice that in generalizing Lemma 2 to Lemma 3, we require a fairly strong condition (for our purposes at least) to hold—it must be the case that the intersection of *any* subset of A_i's is measurable, which is just the thing that we are not requiring when we move from classical probability spaces to generalized probability spaces.

Corollary. If (X, Σ, μ) is a generalized probability space and $A_1, ..., A_n \in \Sigma$, then if for all $B \subseteq \{A_1, ..., A_n\}$, $\bigcap B \in \Sigma$ and $\mu(\bigcap B) = 0$, then $\bigcup_{i=1}^{n} A_i \in \Sigma$ and

$$\mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i)$$

One strange fact follows from the above: we can see that a finite union of measure zero sets in a generalized probability space does not necessarily have measure zero, even if it is measurable. In fact, we can have a finite union of measure zero sets that covers the entire space X, thus receiving measure one. And we can even require that each set in the union be disjoint from some other, and they will still be able to cover the entire space.

Proposition 2. There is a generalized probability space (X, Σ, μ) with sets $A_1, ..., A_n \in \Sigma$ such that:

(i) $\mu(A_i) = 0$, for all $i \leq n$,

(ii) For each $i \leq n$, there is a $j \leq n$ such that $A_i \cap A_j = \varnothing$, and

(iii) $A_1 \cup ... \cup A_n = X$.

Proof. Let $X = \{1, ..., 5\}$, and

$A_1 = \{1, 5\}$

$A_2 = \{2, 5\}$

$A_3 = \{3, 5\}$

$A_4 = \{4\}$

Notice that A_4 is disjoint from A_1, A_2, A_3, but no other pairs are disjoint. Furthermore, $X - A_4 = A_1 \cup A_2 \cup A_3$. Let $\Sigma = \{\varnothing, A_1, A_2, A_3, A_4, X - A_1, X - A_2, X - A_3, A_1 \cup A_2 \cup A_3, A_1 \cup A_4, A_2 \cup A_4, A_3 \cup A_4, X - (A_1 \cup A_4), X - (A_2 \cup A_4), X - (A_3 \cup A_4), X\}$. And we generate μ by additivity from the following assignments:

$\mu(\varnothing) = \mu(A_1) = \mu(A_2) = \mu(A_3) = \mu(A_4) = 0$

$\mu(X) = 1$

One can easily check that μ takes on only the values zero or one for every element of Σ. It follows that (X, Σ, μ) is a generalized probability space that satisfies the constraints. □

The preceding proposition exhibits a very strange feature of generalized probability spaces. Intuitively, we would not expect the disjunction of a finite number of probability zero events to have anything but probability zero. However, in a generalized probability space we do not prohibit the disjunction of a finite number of probability zero events from having probability one, and even covering the entire space. One might consider adding to the axioms of a generalized probability space the following, seemingly weak, condition:

(*) There is not a finite collection of sets $A_1, ..., A_n \in \Sigma$ such that

(i) $\mu(A_i) = 0$, for all $i \leq n$,

(ii) For each $i \leq n$, there is a $j \leq n$ such that A_j is non-empty and $A_i \cap A_j = \varnothing$, and

(iii) $A_1 \cup ... \cup A_n = X$.

While condition (*) may fail as in Proposition 2, there are certain cases where it must hold. If n is 2 or 3, then (*) holds automatically from the axioms of our generalized probability spaces.

Proposition 3. In a generalized probability space (X, Σ, μ), there is no collection of sets $A_1, A_2 \in \Sigma$ such that

(i) $\mu(A_i) = 0$, for all $i \leq 2$,

(ii) $A_1 \cup A_2 = X$.

Proof. Suppose (i) and (ii) hold. Then $(X - A_1) \cap (X - A_2) = \varnothing$ so $\mu((X - A_1) \cup (X - A_2)) = \mu(X - A_1) + \mu(X - A_2) = 1 + 1 = 2$, and $\mu(A_1 \cap A_2) = \mu(X - ((X - A_1) \cup (X - A_2))) = -1$, which contradicts axiom (ii) of generalized probability spaces. □

Proposition 4. In a generalized probability space (X, Σ, μ), there is no collection of sets $A_1, A_2, A_3 \in \Sigma$ such that

(i) $\mu(A_i) = 0$, for all $i \leq 3$,

(ii) $A_1 \cap A_2 = \varnothing$, and

(iii) $A_1 \cup A_2 \cup A_3 = X$.

Proof. Suppose (i), (ii), and (iii) hold. Then $X - A_1 \cup A_2 \in \Sigma$ and is disjoint from $X - A_3$. So $\mu((X - A_1 \cup A_2) \cup (X - A_3)) = 1 + 1 = 2$. It follows that $\mu((A_1 \cup A_2) \cap A_3) = \mu(X - ((X - A_1 \cup A_2) \cup (X - A_3))) = -1$, contradicting axiom (ii) of generalized probability spaces. □

Any generalized probability space that doesn't satisfy (*) exhibits a certain kind of pathology. One might have thought our very concept of probability together with logic requires the disjunction of a finite number of probability zero events to have probability zero[18]. If a generalized probability space fails to satisfy (*), then it breaks this constraint radically. Furthermore, one might think that it's part of the concept of probability that it may serve as a guide to rational action. But any person who sets their subjective degrees of belief by a generalized probability space that violates (*) exhibits a form of irrationality: she will take a series of bets (a Dutch book) for which she is guaranteed to lose money. Suppose we have a sequence of outcomes A_i that violate (*). The generalized probability spacer will sell us \$1 bets on each outcome A_i for free since she assigns probability zero to each of those events, and she will buy a \$1 bet[19] on $\bigcup_{i=1}^{n} A_i = X$ for \$1. She is guaranteed to lose on at least one of these bets, so she will always have a net loss. To rule out these irrationalities, we focus our attention only on generalized probability spaces that satisfy (*).

One might worry that adding condition (*) to the axioms of a generalized probability space might just bring us back to a classical probability space. It turns out it does not.

Proposition 5. There are generalized probability spaces that satisfy (*), which are not also classical probability spaces.

Proof. Consider $X = \{1, 2, 3, 4\}$, and

$A_1 = \{1, 4\}$

$A_2 = \{2, 4\}$

$A_3 = \{3, 4\}$

Notice none of A_1, A_2, A_3 are disjoint. Let $\Sigma = \{\varnothing, A_1, A_2, A_3, (X - A_1), (X - A_2), (X - A_3), X\}$.

[18]While (i) and (iii) seem to have an intuitive justification, (ii) may seem less plausible. The motivation for (ii) is that requiring some pattern of disjointness among our sets should force the additivity axiom to kick in. Also, notice that adding (ii) strictly weakens the condition. We use (ii) in this paper in order to assume only the weakest addition to the axioms for a generalized probability space we can think of to obtain the following results. If one wishes, (ii) can be ignored in order to make the condition more intuitive. Since the resulting condition implies (*), the result of section 5 will still go through on that assumption as well.

[19]This bet on the union is included to show that this is an incoherence in a belief system as opposed to a disagreement between a person's beliefs and the world. If somehow none of the outcomes A_i were to occur, then the generalized probability spacer would not lose on any of the individual A_i bets, but would lose on the union.

We generate μ by additivity from the following assignments:

$\mu(A_1) = \mu(A_2) = \mu(A_3) = 1/3$

$\mu(X) = 1$

It follows that (X, Σ, μ) is a generalized probability space, it satisfies (*) because the only measure zero set is the empty set, and it is not a classical probability space because some intersections of measurable sets are not measurable. $\square$

One might object[20] that the example I provided in the previous proof is, in a sense, not good enough. I have just taken a classical probability space and deleted some of it's assignments, but the space has a *classical probability space extension* as in the following definition.

Definition 9. A generalized probability space (X, Σ, μ) has a *classical probability space extension* iff there is a classical probability space (X, Σ', μ') such that $\Sigma \subseteq \Sigma'$ and if $A \in \Sigma$, then $\mu(A) = \mu'(A)$.

Proposition 6. There are generalized probability spaces that satisfy (*), which have no classical probability space extensions.

Proof. Consider $X = \{1, 2, 3, 4\}$, and

$A_1 = \{1, 4\}$

$A_2 = \{2, 4\}$

$A_3 = \{3, 4\}$

Notice none of A_1, A_2, A_3 are disjoint. Let $\Sigma = \{\varnothing, A_1, A_2, A_3, (X - A_1), (X - A_2), (X - A_3), X\}$. We generate μ by additivity from the following assignments:

$\mu(A_1) = 1/4$

$\mu(A_2) = 0$

$\mu(A_3) = 1/2$

$\mu(X) = 1$

[20]Thanks to Sam Fletcher for this point.

It follows that (X, Σ, μ) is a generalized probability space, it satisfies (*) because the only measure zero sets are the empty set and A_2, but their union is not the whole space. It is not a classical probability space because some intersections of measurable sets are not measurable.

Suppose (X, Σ', μ') were a classical probability space extension of (X, Σ, μ), then $\mu'(A_1 \cup A_2 \cup A_3) = \mu'(X) = 1$, but $\mu'(A_1 \cup A_2 \cup A_3) \leq \mu'(A_1) + \mu'(A_2) + \mu'(A_3) = \mu(A_1) + \mu(A_2) + \mu(A_3) = 3/4$, which is a contradiction. So (X, Σ, μ) is a generalized probability space that satisfies (*) and has no classical probability space extension. □

Notice that the generalized probability space used in the proof of the previous proposition falls prey to another Dutch book. A generalized probability spacer who sets their degrees of belief by the above assignments will sell a \$4 bet on A_1 for \$1, sell a \$4 bet on A_2 for free, sell a \$4 bet on A_3 for \$2, and buy a \$4 bet[21] on $A_1 \cup A_2 \cup A_3 = X$ for \$4. No matter the outcome, she will have a net loss of \$1. This Dutch book arises because the example violated the following condition:

(**) There is not a finite collection of sets $A_1, ..., A_n$ such that

$$\mu(\bigcup_{i=1}^{n} A_i) > \sum_{i=1}^{n} \mu(A_i)$$

Any time condition (**) is violated, one can construct a Dutch book. Of course condition (**) always holds for classical probability spaces—it follows from the classical inclusion-exclusion formula. Notice that condition (**) implies (*). So we have just seen two ways in which (**) does not always hold in generalized probability spaces: in the proof of proposition 2 we exhibited a generalized probability space that fails to satisfy (*) so it fails to satisfy (**), and in the proof of proposition 6 we exhibited a generalized probability space that satisfies (*) but fails to satisfy (**). We also learn from this that (*) does not imply (**).

One might think that, because of these Dutch books, we should only consider generalized probability spaces that satisfy (**). But it is an open question whether there are any generalized

[21] Once again we include this bet on the union to show a strong sense of incoherence. If somehow none of A_1, A_2, or A_3 were to occur, the generalized probability spacer would still have a net loss.

probability spaces satisfying (**), and hence avoiding Dutch books, that do not have classical probability space extensions[22]. If there were not, then we would have reason to reject generalized probability spaces without even considering their compatibility with quantum mechanics. However, since it is an open question, we do not have the resources to rule out generalized probability spaces *a priori*. Instead, in the rest of this paper, we will directly consider the compatibility of generalized probability spaces with quantum mechanics. We will only use condition (*) because it is strictly weaker than (**) and it is all that is needed to prove our main result. In any case, a generalized probability spacer must accept a condition *at least as strong as* (*) to avoid Dutch books.

The above should motivate the consideration of condition (*). We know it always holds in classical probability spaces, and we know it holds in some simple cases ($n = 2$ or 3) for generalized probability spaces. We seem to have good reason to accept a condition at least as strong as (*) in order to avoid Dutch books. However, we now proceed to show that if we add condition (*) to our axioms, then generalized probability spaces cannot solve the problems of quantum mechanics.

5 A "No-Go" Theorem

Definition 10. A quantum mechanical experiment $(\mathcal{H}, \psi, \mathfrak{S})$ has a *restricted generalized probability space representation* iff there is a generalized probability space (X, Σ, μ) with sets $A_1, ...A_n \in \Sigma$ such that for all P_i, P_j,

$$\mu(A_i) = p_i = \langle \psi, P_i \psi \rangle$$

and if $[P_i, P_j] = 0$, then $A_i \cap A_j \in \Sigma$ and

$$\mu(A_i \cap A_j) = p_{ij} = \langle \psi, P_i P_j \psi \rangle$$

[22]We conjecture that all generalized probability spaces satisfying (**) have classical probability space extensions. Here is a sketch of an argument for that claim. Consider a generalized probability space that satisfies (**). Go through all of the pairwise intersections of measurable sets that are not measurable and assign them measures that are consistent with the classical inclusion-exclusion formula. Then go through all of the intersections of three measurable sets and do the same, etc. Eventually all intersections will be assigned a measure and we conjecture the result will be a classical probability space extension of the original space.

Definition 11. A quantum mechanical experiment $(\mathcal{H}, \psi, \mathfrak{S})$ has a *full-blown generalized probability space representation* iff there is a generalized probability space (X, Σ, μ) with sets $A_1, ..., A_n \in \Sigma$ such that for all P_i,

$$\mu(A_i) = p_i = \langle \psi, P_i \psi \rangle$$

and for all $P_i, P_j, ..., P_k$ that are compatible (i.e. they pairwise commute), the corresponding sets $A_i, A_j, ..., A_k$ satisfy $A_i \cap A_j \cap ... \cap A_k \in \Sigma$ and

$$\mu(A_i \cap A_j \cap ... \cap A_k) = p_{ij...k} = \langle \psi, P_i P_j ... P_k \psi \rangle$$

Notice that since all classical probability spaces are generalized probability spaces, the result of Theorem 4 carries over. If all of the observables of a experimental system commute, then it has a restricted generalized probability space representation. But now one can ask the following question: is there a generalized probability space representation for every quantum mechanical experiment?

In order to answer this question, we'll use another well-known "no-go" theorem, due to Kochen and Specker, concerning HVTs in quantum mechanics. An alternative way to construct a HVT would have been to construct a function whose inputs are observables and outputs are definite real numbers—we would then require that this function satisfies certain natural conditions for exhibiting the complete hidden state. Kochen and Specker showed that if we require our function to yield consistent answers to every "yes-no" question, then we are guaranteed that there is a quantum mechanical experiment for which we will not be able to find this kind of HVT.

Theorem 5. (Kochen-Specker) For any Hilbert space $\mathcal{H}$ with $dim(\mathcal{H}) \geq 3$, there is a sequence of projection operators $\mathfrak{S}' = \langle P_1, ..., P_n \rangle$ on $\mathcal{H}$ such that there is no function $f : \{P_1, ..., P_n\} \rightarrow \{0, 1\}$ which assigns 1 to exactly one element of every subset of $\{P_1, ..., P_n\}$ whose elements are mutually orthogonal and span $\mathcal{H}$ (Kochen and Specker 1967, p. 321).

Now, we return to the main question of this section. One can show, although I won't do so here, that some quantum mechanical experiments[23] do have (restricted and full-blown) generalized probability space representations, and even ones which satisfy (*). This motivates the question: do generalized probability spaces provide a way of finding HVTs for *all* quantum mechanical experiments?

Theorem 6. It is not the case that all quantum mechanical experiments $(\mathcal{H}, \psi, \mathfrak{S})$ have full-blown[24] generalized probability space representations (X, Σ, μ) that satisfy condition (*).

Proof.[25] Suppose we have an $\mathcal{H}$ and sequence of projection operators $\mathfrak{S}' = \langle P_1, ..., P_n \rangle$ on $\mathcal{H}$ that provide a witness to Theorem 5 i.e. there is no $f : \{P_1, ..., P_n\} \to \{0, 1\}$ which assigns 1 to exactly one element of every subset of $\mathfrak{S}'$ whose elements are mutually orthogonal and span $\mathcal{H}$.

Fix some unit vector $\psi \in \mathcal{H}$, and let $\mathfrak{S} = \langle P_1, ..., P_n, (\mathbb{I} - P_1), ..., (\mathbb{I} - P_n) \rangle = \langle P_1, ..., P_{2n} \rangle$, which is still a finite set of projection operators. Now suppose (for contradiction) that for our chosen quantum mechanical experiment $(\mathcal{H}, \psi, \mathfrak{S})$, we have the desired full-blown generalized probability space (X, Σ, μ) that satisfies (*). On this assumption, we construct a function $f : \{P_1, ..., P_{2n}\} \to \{0, 1\}$, which when restricted to the domain $\{P_1, ..., P_n\} \subseteq \{P_1, ..., P_{2n}\}$ contradicts Theorem 5.

We would like to fix some point $x \in X$ and determine the values of f just by whether x falls in the set corresponding to a given projection operator. But x may end up in a measure zero set such that it is in none of the sets corresponding to projection operators that span the space, or such that it is in the intersection of two sets corresponding to orthogonal projection operators. If we end up in either of these cases, then the function f we construct will not violate the Kochen-Specker theorem. So we construct two "problematic" sets to remove from our space—these contain the superfluous measure zero sets whose removal will not make a difference to the physics. Removing these sets guarantees we will be able to violate the Kochen-Specker theorem in the way just outlined.

[23]In particular, see Malament (2012, p. 27) for a proof that the quantum mechanical experiment representing the EPR setup, which served as a witness to Bell's theorem and Pitowsky's theorem, has a generalized probability space representation satisfying (*). The generalized probability space exhibited there does not, however, satisfy (**).

[24]It remains an open question whether the result holds for restricted generalized probability space representations. It also remains an open question whether the result holds if we weaken or get rid of condition (*).

[25]Thanks to David Malament for detailed discussions concerning the proof of this theorem.

We know that for any orthogonal P_i and P_j (written $P_i \perp P_j$), $P_i P_j = P_j P_i = 0$ so $A_i \cap A_j \in \Sigma$ and $\mu(A_i \cap A_j) = \langle \psi, P_i P_j \psi \rangle = 0$. So, let

$$D_1 = \{A_i \cap A_j : P_i \perp P_j\}$$
$$R_1 = \bigcup D_1$$

Each element of D_1 is a set of hidden states that has two "contradictory" properties, i.e. properties corresponding to orthogonal projection operators. So D_1 is a finite collection of measure zero sets.

Next, for any $Q \subseteq \{P_1, ..., P_{2n}\}$ whose members are mutually orthogonal and span $\mathcal{H}$, let $R_Q = \bigcup_{P_i \in Q} A_i$. Let

$$D_2 = \{X - R_Q : \text{ the members of } Q \text{ are mutually orthogonal and span } \mathcal{H}\}$$
$$R_2 = \bigcup D_2$$

Each element of D_2 is a set of hidden states that do not have a "tautologous" property, i.e. the property corresponding to lying in at least one of a set of mutually orthogonal subspaces that span $\mathcal{H}$. We expect D_2 to also be a finite collection of measure zero sets; we show this explicitly below.

Let $X' = X - R_1 \cup R_2$ and $A'_i = A_i \cap X'$.

We now check that X' is non-empty so we can use a point in it to define our function f, and that the sets A'_i behave as expected, i.e. they allow us to construct the required function f.

Subclaim 1: X' is non-empty.

We show that $R_1 \cup R_2$, the set to be removed, is a finite collection of measure zero sets, each of which is disjoint from at least one other in the collection. In order for the space to satisfy (*), there must be points in X that are not in $R_1 \cup R_2$.

For any $Q \subseteq \{P_1, ..., P_{2n}\}$ whose members are mutually orthogonal and span $\mathcal{H}$, we know that for any set of projection operators $\{P_i, P_j, ..., P_k\} \subseteq Q$, the corresponding sets satisfy[26] $A_i \cap A_j \cap ... \cap A_k \in \Sigma$ and $\mu(A_i \cap A_j \cap ... \cap A_k) = \langle \psi, P_i P_j ... P_k \psi \rangle = 0$. Thus, by the Corollary to

[26]Notice that this is the only place in the proof where we use the fact that we are dealing with a full-blown, rather than restricted representation.

Lemma 3, $R_Q \in \Sigma$ and $\mu(R_Q) = \sum_{P_i \in Q} \mu(A_i) = \sum_{P_i \in Q} \langle \psi, P_i \psi \rangle = 1$. Hence, $X - R_Q \in \Sigma$ and $\mu((X - R_Q) \cup R_Q) = \mu(X - R_Q) + \mu(R_Q) = \mu(X)$. It follows that $\mu(X - R_Q) + 1 = 1$, so $\mu(X - R_Q) = 0$.

Now we have that R_1 is a finite union of sets all of whose measure is zero, and R_2 is a finite union of sets all of whose measure is zero. So $R_1 \cup R_2 = \bigcup(D_1 \cup D_2)$ is likewise a finite union of sets all of whose measure is zero, i.e. for any set $B \in (D_1 \cup D_2)$, $\mu(B) = 0$, and $D_1 \cup D_2$ is a finite collection of subsets of X.

Furthermore, if $P_i \perp P_j$, then there is some Q whose members are mutually orthogonal and span $\mathcal{H}$ such that Q contains P_i (at the very least P_i and $\mathbb{I} - P_i$ are orthogonal and span $\mathcal{H}$). Moreover, for this choice of Q, $A_i \cap A_j$ and $X - R_Q$ are disjoint, where A_i and A_j are the sets associated with P_i and P_j and R_Q is defined relative to the Q containing P_i. Thus, for each $B \in D_1$, there is a $C \in D_2$ such that $B \cap C = \varnothing$.

And given any Q whose members are mutually orthogonal and span $\mathcal{H}$, we can choose two of the orthogonal projection operators $P_i, P_j \in Q$. For this choice, $A_i \cap A_j$ and $X - R_Q$ are disjoint, where again A_i and A_j are the sets associated with P_i and P_j, and R_Q is defined relative to the Q containing P_i and P_j. Thus, for every $C \in D_2$, there is a $B \in D_1$ such that $B \cap C = \varnothing$.

So $D_1 \cup D_2$ is a finite collection of measure zero sets, each of which is disjoint from at least one of the others. By (*), $R_1 \cup R_2 = \bigcup(D_1 \cup D_2) \neq X$. Therefore, X' is non-empty.

Subclaim 2: If the members of $Q \subseteq \{P_1, ..., P_{2n}\}$ are mutually orthogonal and span $\mathcal{H}$, then $\bigcup_{P_i \in Q} A'_i = X'$.

If the members of $Q \subseteq \{P_1, ..., P_{2n}\}$ are mutually orthogonal and span $\mathcal{H}$, then $(X - R_Q) \subseteq R_2$, so since $X' \cap R_2 = \varnothing$, $X' \cap (X - R_Q) = \varnothing$. But since $X' \subseteq X$, $X' \cap (X - R_Q) = X' - X' \cap R_Q = \varnothing$. Hence, $X' \cap R_Q = X'$. Thus, $\bigcup_{P_i \in Q} A'_i = \bigcup_{P_i \in Q}(A_i \cap X') = X' \cap \bigcup_{P_i \in Q} A_i = X' \cap R_Q = X'$.

Subclaim 3: If $P_i \perp P_j$, then $A'_i \cap A'_j = \varnothing$.

If $P_i \perp P_j$, then $(A_i \cap A_j) \subseteq R_1$ and $R_1 \cap X' = \varnothing$ so $(A_i \cap A_j) \cap X' = \varnothing$. Thus, $A'_i \cap A'_j = (A_i \cap X') \cap (A_j \cap X') = (A_i \cap A_j) \cap X' = \varnothing$.

Fix some $x \in X'$. Let $f : \{P_1, ..., P_{2n}\} \to \{0, 1\}$ be defined for any P_i by:

$$f(P_i) = \chi_{A_i'}(x),$$

where $\chi_{A_i'}$ is the characteristic function of A_i' in X' (i.e. $\chi_{A_i'}(x) = 1$ if $x \in A_i'$ and $\chi_{A_i'}(x) = 0$ if $x \notin A_i'$).

Consider any $Q \subseteq \{P_1, ..., P_{2n}\}$ whose elements are mutually orthogonal and span $\mathcal{H}$. Since the members of Q span $\mathcal{H}$, we know that $x \in A_i'$ for at least one $P_i \in Q$ by subclaim 2. And since all members of Q are orthogonal, we know that $x \in A_i'$ for at most one $P_i \in Q$ by subclaim 3. Thus, $x \in A_i'$ for exactly one $P_i \in Q$, and it follows that $f(P_i) = 1$ for exactly this one member of Q.

Thus, on the assumption that our desired full-blown generalized probability space representation exists, we can construct the function f that is prohibited by the Kochen-Specker theorem. From this we conclude that the desired generalized probability space representation does not exist. □

6 Conclusions

We have seen that the consideration of incompatible observables is well motivated in an investigation of the foundations of probability in quantum mechanics. We can always find a HVT in the form of a classical probability space if all of the observables of our experimental system are compatible, so something funny must be going on when we try to represent incompatible observables in a classical probability space. The proposal to use generalized probability spaces seemed like a natural alternative which might solve these problems. In at least a few places in the literature, the use of generalized probability spaces has been suggested for understanding quantum mechanics.

However, upon investigation, we found that generalized probability spaces have a very strange feature, namely allowing the union of finitely many probability zero events to cover the whole space. One might have thought that it was a fundamental property of probability spaces, coming from the very concept of probability itself, that the union of finitely many probability zero events cannot cover the whole space. We do find this to be the case in classical probability spaces. One may have thought it a virtue of classical probability spaces that they preserved this property with only a few simple axioms, and one might have thought that the only reason we don't add condition (*) as an extra axiom in classical probability spaces is that it is implied by the axioms we have already displayed. Furthermore, if one sets their degrees of belief by a generalized probability space that fails to satisfy (*), then one is susceptible to a Dutch book. So one might think that we should add condition (*) to the axioms for generalized probability spaces, since it is not implied by them.

When we added condition (*) to the axioms of generalized probability spaces, we found that one cannot represent all quantum mechanical experiments in this way. We cannot find a HVT in the form of a classical probability space for all quantum mechanical experiments nor can we find a HVT in the form of a generalized probability space that satisfies (*) for all quantum mechanical experiments. If one wants to find HVTs in the form of generalized probability spaces for all quantum mechanical experiments, then one must pay the price of giving up condition (*). The generalized probability spacer is impaled on the horns of a dilemma: either give up (*) and fall prey to Dutch books, or else accept (*) and run into Theorem 6. This shows a sense in which the proposal of getting rid of joint distributions on non-commuting observables and using generalized probability spaces as HVTs for quantum mechanics fails.

An interesting corollary follows immediately from Theorem 6. Since the Kochen-Specker theorem (Theorem 5) is used essentially in the proof of Theorem 6, it follows that the Kochen-Specker theorem implies Theorem 6. And since all classical probability spaces are generalized probability spaces satisfying (*), it also follows that Theorem 6 implies Pitowsky's theorem (Theorem 3). So the Kochen-Specker theorem also implies Pitowsky's theorem. And furthermore, since it was shown in section 2.2 that Pitowsky's theorem implies Bell's theorem, this means that the Kochen-Specker theorem implies Bell's theorem as well. This establishes a hierarchy of implication relations between the "no-go" theorems:

$$\text{Kochen-Specker Theorem} \Rightarrow \text{Theorem 6} \Rightarrow \text{Pitowsky's Theorem} \Rightarrow \text{Bell's Theorem}$$

One might have thought initially that Bell's theorem, Pitowsky's theorem, and the Kochen-Specker theorem were three independent results that show we cannot have HVTs with different sorts of constraints on them. While Bell's theorem and Pitowsky's theorem deal with HVTs that represent the probabilistic predictions of quantum mechanics, the Kochen-Specker theorem puts only purely logical constraints on an HVT without any mention of probabilistic predictions. Now we can see that even though these theorems deal with different constraints on HVTs, they are systematically related. Just as in section 2.2 we thought of Bell's theorem as a specialization of Pitowsky's theorem to situations where we have further information about distinct measurement apparatuses and measurement settings, we can think of Theorem 6 and Pitowsky's theorem as specializations of the Kochen-Specker theorem to situations in which we have further information about probabilistic predictions (in the relevant kind of probability space) and maintain the natural logical constraints. This is a promising beginning to the project of mapping out a conceptual space in which to compare the constraints and results of these "no-go" theorems.

Acknowledgements

Thanks to Jim Weatherall, Jeff Barrett, Sam Fletcher, Dan Shiff, and the Southern California Philosophy of Physics Reading Group (Oct. 2012) for helpful comments and discussions. I am particularly grateful for detailed comments from David Malament on this paper and the proofs contained in it.

References

[1] Bell, John. "On the Einsten Podolsky Rosen Paradox." *Physics*. 1.3 (1964): 195-200.

[2] Billingsley, Patrick. *Probability and Measure*. NY: John Wiley & Sons, Inc., 1979.

[3] Bohm, David and Yakir Aharanov. "Discussion of Experimental Proof for the Paradox of Einstein, Rosen, and Podolsky" *Physical Review*. 108.4 (1957): 1070-1076.

[4] Einstein, Albert, B. Podolsky, and N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" *Physical Review*. 47 (1935): 777-780.

[5] Fine, Arthur. "Joint distributions, quantum correlations, and commuting observables." *Journal of Mathematical Physics*. 23.7 (1982a): 1306-1310.

[6] Fine, Arthur. "Hidden Variables, Joint Probability, and the Bell inequalities." *Physical Review Letters*. 48.5 (1982b): 291-294.

[7] Gudder, Stanley. *Quantum Probability*. San Diego, CA: Academic Press, Inc., 1988.

[8] Hooker, C. (ed.). *The Logico-Algebraic Approach to Quantum Mechanics*. Dordrecht: Reidel, 1975.

[9] Kochen, S. and E. Specker. "The Problem of Hidden Variables in Quantum Mechanics." *Journal of Mathematics and Mechanics*. 17 (1967): 5987; reprinted in Hooker 1975, 293-328 (page references to reprint).

[10] Krantz, David, Duncan Luce, Patrick Suppes, and Amos Tversky. *Foundations of Measurement*, vol. I. Mineola, NY: Dover Publications, Inc., 1971.

[11] Malament, David. "Notes on Bell's Theorem." Unpublished Lecture Notes, accessed 2012. http://www.socsci.uci.edu/ dmalamen/courses/prob-determ/PDnotesBell.pdf

[12] Pitowsky, Itamar. *Quantum Probability—Quantum Logic*. New York: Springer-Verlag, 1989.

[13] Suppes, Patrick. "The Probabilistic Argument for a Nonclassical Logic of Quantum Mechanics." *Philosophy of Science*. 33 (1966): 14-21; reprinted in Hooker 1975, 341-350 (page references to reprint).

[14] Suppes, Patrick. "Logics Appropriate to Empirical Theories." *Symposium on the Theory of Models*. Amsterdam: North Holland Publishing Company, 1965; reprinted in Hooker 1974, 329-340 (page references to reprint).