# Hidden Variables and Commutativity in Quantum Mechanics*

Benjamin Feintzeig‡

October 10, 2012

## 1 Introduction

Various well known "no-go" theorems purport to show that one cannot find hidden variable theories (HVTs) for all quantum mechanical systems. Some have suggested that the reason we cannot find HVTs is that we expect too much of them. The usual way of constructing a HVT requires that we assign 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 that does just this—while HVTs are typically constructed as *classical probability spaces*[1], 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

*Draft - PLEASE DO NOT QUOTE OR CITE!

†*Email*: bfeintze@uci.edu

‡Thanks to David Malament, Jim Weatherall, Jeff Barrett, Sam Fletcher, and Dan Shiff for helpful comments and discussions concerning this paper.

[1]Some might not have realized that classical probability spaces, as defined in section 2.1, are the usual way of constructing HVTs. For this reason, I show the relationship between classical probability spaces and perhaps more typical ways of thinking about HVTs in that same section.

systems 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. Then in section 5, we show just what it would take for there to be generalized probability space representations for all quantum mechanical systems.

## 2 Bell's Theorem and Classical Probability Spaces

**Definition 1.** For the purposes of this paper, we take a *quantum mechanical system* 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[2] onto subspaces of $\mathcal{H}$.

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:

$$p_{ij} = \langle\psi, P_iP_j\psi\rangle$$

A quantum mechanical system brings with it a data set of type $(n, S)$, where $n$ is the number of projection operators in $\mathfrak{S}$, and $S = \{\langle i, j\rangle : [P_i, P_j] = 0\}$. Using the above, 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.

[2]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

First, we give a typical presentation of Bell's theorem before moving on to Pitowsky's slightly more abstract variation on Bell's theorem, which considers classical probability spaces.

The usual understanding of Bell's theorem takes the settings of our measurement apparatus into account in addition to the measurement outcomes. 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 $i$th measurement apparatus given that it was prepared with measurement setting $a_i$. For example, if we consider the EPR setup[3], 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 $i = 2$, 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. We determine these probabilities from the Hilbert space associated with the system using the above prescription.

**Definition 2.** A *Bell HVT* for a quantum mechanical system 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 $\rho$ is a probability density (i.e. $\rho : X \to [0, 1]$ and $\int_X \rho(x)dx = 1$) such that[4]

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

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

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

[4] 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.

the measurement settings, but also on the hidden state $x$. The probability density $\rho(x)$ represents the probability of finding the system in the hidden state $x$.

Now we put constraints on the the probability functions of the HVT. Although there are many ways to do it, for our purposes only two constraints are relevant. We formulate the constraints in terms of the EPR setup in which $n = 2$.

**Quasi-determinism:** 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 $A_1, A_2, a_1, a_1', a_2, a_2', x$,

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

In a Quasi-deterministic HVT, the system has a definite value for each one of it's properties: either "yes" if the probability is 1 or "no" if the probability is zero. In a Local HVT, 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-determinism and Locality, one can derive a characteristic inequality, known as Bell's inequality. So if all quantum mechanical systems had Local, Quasi-deterministic HVTs, then the probabilistic predictions of each of those quantum mechanical systems would satisfy Bell's inequality. But when we consider the quantum mechanical system 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 system violates Bell's inequality (Bell 1964, p. 198; Pitowsky 1989, p. 84). So we conclude that there could not be a Local, Quasi-deterministic hidden variable theory for all quantum mechanical systems—in particular, not for the EPR setup.

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

## 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 *$\sigma$-algebra* $\Sigma$ 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)[5] 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 $(X, \Sigma, \mu)$, where $X$ is a non-empty set of states, $\Sigma$ is a $\sigma$-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 system, and each measurable set in $\Sigma$ corresponds to a property that the system can either have or not have by falling in or out of that set. If we knew the hidden state of our system, we would know all of its properties, but since we don't, we can only ascribe probability values to the properties of our 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

[5] We restrict our attention in this paper to finite unions although in general, a $\sigma$-algebra allows for countable unions, and the axioms of the probability space change accordingly. Technically, what I call $\sigma$-algebras here are mere algebras, and what I call $\sigma$-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 5.

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

**Definition 5.** A quantum mechanical system $(\mathcal{H}, \psi, \mathfrak{S})$ has a *restricted classical probability space representation* iff there is a classical probability space $(X, \Sigma, \mu)$ 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 4, 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 system $(\mathcal{H}, \psi, \mathfrak{S})$ has a *full-blown classical probability space representation* iff there is a classical probability space $(X, \Sigma, \mu)$ 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-deterministic Bell HVT leads to a classical probability space representation in the following way.

**Proposition 1.** If $(X, \{pr_{HV}(A_1, A_2|a_1, a_2; x) : x \in X\}, \rho)$ constitutes a Local, Quasi-deterministic Bell HVT for a specified quantum mechanical system, then there is a classical probability space $(X, \Sigma, \mu)$ such that for all $a_1, a_2$, there are sets[6] $L_{a_1}, R_{a_2}, \in \Sigma$ for which[7]

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

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

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

[6] $L_{a_1}$ 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 left, given setting $a_2$.

[7] I.e. the specified quantum mechanical system 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 $\Sigma$ be the $\sigma$-algebra generated by all of the sets of the form $L_{a_1}$ and $R_{a_2}$. Define $\mu$ 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 $(X, \Sigma, \mu)$ 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-deterministic Bell HVT comes with a classical probability space representation, showing that there are quantum mechanical systems with no classical probability space representation would imply that there are quantum mechanical systems with no Local, Quasi-deterministic Bell HVT. We now turn our attention to whether there are classical probability space representations for all quantum mechanical systems.

Given a quantum mechanical system 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 $p^\epsilon$, 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[8] 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 systems violate these inequalities. So we conclude with a "no-go" theorem for classical probability space representations as HVTs. Notice that the following implies Bell's theorem (Theorem 1) via Proposition 1.

[8]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 10 for more.

**Theorem 3.** (Pitowsky/Bell) There are quantum mechanical systems for which there are no restricted classical probability space representations, and therefore there are quantum mechanical systems for which there are no full-blown classical probability space representations.

## 3 Incompatible Observables

Arthur Fine argues that the reason we cannot find HVTs in the form of classical probability spaces for quantum mechanical systems is that we are trying to assign probability values 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 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 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 concerning joint distributions and compatible observables[9] to motivate these claims. Here, we present yet another variant, using Pitowsky's powerful results.

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

**Theorem 4.** (Fine) For all quantum mechanical systems $(\mathcal{H}, \psi, \mathfrak{S})$, 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 system has a restricted[10] classical probability space representation.

*Proof.*[11] 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 $p_{ij} = \langle \psi, P_i P_j \psi\rangle = \sum_{\{\epsilon \in I : \epsilon_i = \epsilon_j = 1\}} \langle \psi, P(\epsilon)\psi\rangle = \sum_{\{\epsilon \in I : \epsilon_i = \epsilon_j = 1\}} \lambda(\epsilon) = \sum_{\epsilon \in I} \lambda(\epsilon)\epsilon_i \epsilon_j = \sum_{\epsilon \in I} \lambda(\epsilon)(p^\epsilon)_{ij}$.

Thus, $\langle p_1, ..., p_n, ..., p_{ij}, ...\rangle \in c(n, S)$, so by Theorem 1, the desired classical probability space representation exists. □

Since we have no trouble finding HVTs, in the form of classical probability spaces when our system involves only compatible observables, one might claim that the reason we fail to find classical probability space represenations for quantum mechanical systems 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.

[10] If one generalized Pitowsky's result in Theorem 2, as we conjectured in footnote 8 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.

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

# 4 Generalized Probability Spaces

**Definition 7.** A *σ-additive class* (Gudder 1988, 90) $\Sigma$ 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 $A \cup B \in \Sigma$.

The only difference between classical and generalized probability spaces comes in condition (ii). We only require unions, and consequently intersections, to be measurable if they are between disjoint sets. The rest of the definition proceeds exactly as for classical probability spaces.

**Definition 8.** A *generalized probability space* (Krantz et al. 1971, p. 214; Gudder 1988, p. 169) is an ordered triple $(X, \Sigma, \mu)$, where $X$ is a non-empty set of states, $\Sigma$ is a $\sigma$-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? We can see immediately from the definition that 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 systems, 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][12] probability spaces can be used as the basis for an axiomatic development of classical quantum mechanics[13]" (Suppes 1966, p. 345-347).

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

[13] 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.

And Gudder writes,

> "[M]uch of quantum mechanics can be described in the framework of $\sigma$-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 systems.

One strange fact about generalized probability spaces is that since unions and intersections of measurable sets are not always required to be measurable, the inclusion-exclusion formula does not hold in general. If $A_1, ..., A_n \in \Sigma$, we are not even guaranteed that $A_1 \cup ... \cup A_n$ is measurable. However, one can prove some slightly weaker results.

**Lemma 1**: If $(X, \Sigma, \mu)$ 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) $C \in \Sigma$.

*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 $(X, \Sigma, \mu)$ is a generalized probability space and $A$,$B \in \Sigma$, then if either $A \cap B$ or $A \cup B$ is in $\Sigma$, 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 $\Sigma$, 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 $(X, \Sigma, \mu)$ 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 $(X, \Sigma, \mu)$ 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 $(X, \Sigma, \mu)$ 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 $\mu$ 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 $\mu$ takes on only the values zero or one for every element of $\Sigma$. It follows that $(X, \Sigma, \mu)$ 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 (i) and (iii) seem to have an intuitive justification, namely that we would not expect the union of finitely many probability zero events to cover the entire space, (ii) may seem less plausible. Notice, however, 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 that we can think of to obtain the following results.[14]

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

**Proposition 3.** 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 $\Sigma = \{\varnothing, A_1, A_2, A_3, (X - A_1), (X - A_2), (X - A_3), X\}$. We generate $\mu$ by additivity from the following assignments:

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

$\mu(X) = 1$

It follows that $(X, \Sigma, \mu)$ 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. □

One might object[15] 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.

[14]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.

[15]Thanks to Sam Fletcher for this point.

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

We can quell the worry with the following proposition.

**Proposition 4.** There are generalized probability spaces that satisfy (*), which are 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 $\mu$ by additivity from the following assignments:

$\mu(A_1) = 1/4$

$\mu(A_2) = 0$

$\mu(A_3) = 1/2$

$\mu(X) = 1$

It follows that $(X, \Sigma, \mu)$ 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 $(X, \Sigma', \mu')$ were a classical probability space extension of $(X, \Sigma, \mu)$, 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 $(X, \Sigma, \mu)$ is a generalized probability space that satisfies (*) and has no classical probability space extension. □

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 5.** In a generalized probability space $(X, \Sigma, \mu)$, 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 6.** In a generalized probability space $(X, \Sigma, \mu)$, 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 $\mu((X - A_1 \cup A_2) \cup (X - A_3)) = 1 + 1 = 2$, and 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. □

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, but we also know it fails in some generalized probability spaces. It should also be somewhat intuitive that the union of finitely many probability zero sets should not be our entire space, especially if those original sets have some pattern of disjointness, since disjointness is what forces the additivity axiom to kick in. 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 that we would like them to.

## 5 A "No-Go" Theorem

We may define generalized probability space representations as we did above for classical probability spaces.

**Definition 10.** A quantum mechanical system $(\mathcal{H}, \psi, \mathfrak{S})$ has a *restricted generalized probability space representation* iff there is a generalized probability space $(X, \Sigma, \mu)$ 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

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

**Definition 11.** A quantum mechanical system $(\mathcal{H}, \psi, \mathfrak{S})$ has a *full-blown generalized probability space representation* iff there is a generalized probability space $(X, \Sigma, \mu)$ 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 quantum 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 system?

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 and to 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 system 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 $f : \{P_1, ..., P_n\} \to \{0, 1\}$ 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: do generalized probability spaces provide a way of finding HVTs for all quantum mechanical systems?

**Theorem 6.**[16] It is not the case that all quantum mechanical systems $(\mathcal{H}, \psi, \mathfrak{S})$ have full-blown[17] generalized probability space representations $(X, \Sigma, \mu)$ that satisfy condition (*).

*Proof.*[18] 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 system $(\mathcal{H}, \psi, \mathfrak{S})$, we have the desired full-blown generalized probability space

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

[17] It remains an open question whether the result holds for restricted generalized probability space representations.

[18] This proof also shows a sense in which the Kochen-Specker theorem (Theorem 5) implies Pitowsky's theorem (Theorem 3), since all classical probability spaces are generalized probability spaces satisfying (*).

$(X, \Sigma, \mu)$ 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 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 $\mu(A_i \cap A_j) = \langle \psi, P_i P_j \psi \rangle = 0$. We construct two "problematic" sets to remove from our space.

First, let

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

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

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

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

*Subclaim 1*: $X'$ is non-empty.

For any $Q \subseteq \{P_1, ..., P_{2n}\}$ whose members are mutually orthogonal and span $\mathcal{H}$, we know that for all $B = \{P_i, P_j, ..., P_k\} \subseteq Q$, $A_i \cap A_j \cap ... \cap A_k \in \Sigma$ 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 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, we know that if $P_i \perp P_j$, then there is some $Q$ containing $P_i$ (at the very least $P_i$ and $\mathbb{I} - P_i$ are orthogonal and span $\mathcal{H}$) so $A_i \cap A_j$ and $X - R_Q$ are disjoint. On the other hand, for every $Q$, we can choose two of the orthogonal projection operators $P_i, P_j \in Q$

so we know $A_i \cap A_j$ and $X - R_Q$ are disjoint. Thus, for each $B \in D_1$, there is a $C \in D_2$ such that $B \cap C = \varnothing$, and for every $C \in D_2$, there is a $B \in D_1$ such that $B \cap C = \varnothing$. $D_1 \cup D_2$ 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 $f : \{P_1, ..., P_{2n}\} \rightarrow \{0, 1\}$ 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 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. $\square$

## 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 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. If one thinks this is a feature that all probability spaces should have in order to represent something resembling our usual concept of probability, then 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 systems in this way. This suggests a sense in which the proposal of getting rid of joint distributions on non-commuting observables fails. We cannot find a HVT in the form of a classical probability space for all quantum mechanical systems nor can we find a HVT in the form of a generalized probability space that satisfies (*) for all quantum mechanical systems. If one wants to find HVTs in the form of generalized probability spaces for all quantum mechanical systems, then one must pay the price of giving up condition (*). While this result does not completely rule out the consideration of generalized probability spaces, it displays some of their unintuitive features, and shows just what it would take for them to play the role we initially hoped for in quantum mechanics.

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

[12] Putnam, Hilary. “Is Logic Empirical?” *Boston Studies in Philosophy of Science V*. Dordrecht, Holland: D. Reidel Publishing Company, 1969; reprinted in Hooker 1975, 181-206 (page references to reprint).

[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).