A classification of hidden-variable properties

2008 J. Phys. A: Math. Theor. 41 425302

(http://iopscience.iop.org/1751-8121/41/42/425302)

IOP Publishing — Journal of Physics A: Mathematical and Theoretical

J. Phys. A: Math. Theor. **41** (2008) 425302 (21pp) doi:10.1088/1751-8113/41/42/425302

A classification of hidden-variable properties

Adam Brandenburger[1] **and Noson Yanofsky**[2]

[1] Stern School of Business, New York University, 44 West Fourth Street, New York, NY 10012, USA

[2] Department of Computer and Information Science, Brooklyn College, 2900 Bedford Avenue, Brooklyn, NY 11210, USA

E-mail: adam.brandenburger@stern.nyu.edu and noson@sci.brooklyn.cuny.edu

Received 8 January 2008, in final form 28 August 2008
Published 26 September 2008
Online at stacks.iop.org/JPhysA/41/425302

Abstract

Hidden variables are extra components added to try to banish counterintuitive features of quantum mechanics. We start with a quantum-mechanical model and describe various properties that can be asked of a hidden-variable model. We present six such properties and a Venn diagram of how they are related. With two existence theorems and three no-go theorems (EPR, Bell and Kochen–Specker), we show which properties of empirically equivalent hidden-variable models are possible and which are not. Formally, our treatment relies only on classical probability models, and physical phenomena are used only to motivate which models to choose.

PACS numbers: 03.65.Ta, 03.65.Ud, 02.50.Cw

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Begun by von Neumann [35, 1932], the hidden-variable program in quantum mechanics (QM) adds extra 'hidden' ingredients in order to try to banish some of the counterintuitive features of QM. These features are: (i) the probabilistic nature of quantum behavior, (ii) the possibility of so-called non-local effects between widely separated particles and (iii) the idea of an intrinsic dependence between the observer of a QM system and the properties of the system itself.

Hidden-variable theories aim to remove these strange aspects of QM by building more 'complete' models (in the terminology of Einstein–Podolsky–Rosen [16, 1935]). The completed models should agree with the predictions of QM, but exhibit one or more of the desired properties of: (i) determinism, (ii) locality and (iii) independence.

Can such models actually be built? The famous 'no-go' theorems of QM show that there are severe limitations to what can be done. But it is also true that certain combinations of properties are possible.

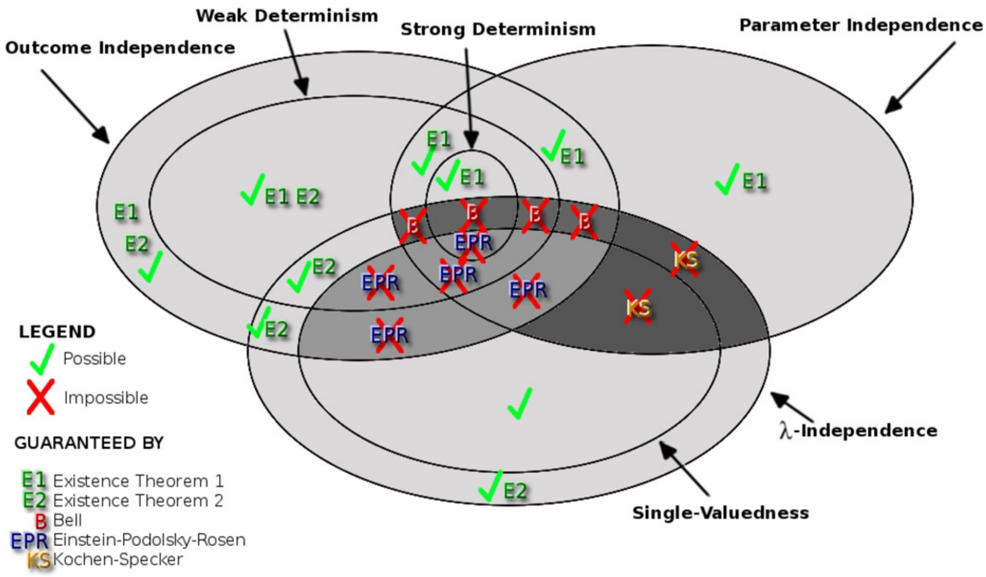

Figure 1. Complete Venn diagram.

Our modest goal in this paper is to provide a formal framework in which various properties one might ask of hidden-variable models can be stated and in which various non-existence and existence results can be organized. Almost all–if not all–of the ingredients of what we do in this paper are well known to researchers in the area. Our contribution, we hope, is in putting all the ingredients into one simple setting.

The setting is classical probability spaces. The question is, given a classical probability model, whether there exists an associated hidden-variable model that is empirically equivalent to the first model and that satisfies certain properties. These properties are motivated by the literature on hidden variables in QM. The specific properties we consider—and the relationships among them—can be depicted in the Venn diagram of figure 1. (We define all the terms later.) The diagram contains 21 regions.

The main result of this paper is that we can give a complete account of these 21 regions. For 10 of these regions (indicated with checks), it is always possible to find an equivalent hidden-variable model with the properties in question. For the remaining 11 regions (indicated with crosses), this may not be possible. We fill in the regions via two existence results and three non-existence results. The latter three are the famous theorems of Einstein–Podolsky–Rosen (EPR) [16, 1935],[3] Bell [2, 1964] and Kochen–Specker [24, 1967].

It is important to understand that, formally, our paper makes no use of physical phenomena. It is an exercise in classical probability theory alone. Of course, the probability spaces we select for the non-existence results are inspired by the physical experiments described in EPR, Bell and Kocher–Specker. But we hope it is conceptually clarifying to present the hidden-variable question in a purely abstract setting—that is, to show how much follows from the rules of probability theory alone.

Naturally, our account of hidden-variable theory is complete only relative to the properties we consider (there are six of them). These properties are, as far as we can tell, the main ones considered in the literature. (As we explain later, we have added one definition.) In

[3] Strictly speaking, EPR did not state a non-existence theorem, but it is useful to present their argument this way.

particular, Bell locality ([2, 1964]) is equivalent to the conjunction of outcome and parameter independence (Jarrett [23, 1984], stated here as proposition 2.1). Kochen–Specker [24, 1967] non-contextuality is implied by the conjunction of parameter independence and λ-independence (proposition 2.2). But there may well be other interesting properties to put on hidden-variable models, which would lead to an extension of figure 1.

There are, of course, many treatments of the hidden-variable question. (Prominent examples include Belinfante [1, 1973], Gudder [21, 1988], Mermin [27, 1993], Peres [29, 1990] and van Fraassen [34, 1991].) This paper is not meant to be a comprehensive survey. Our goal is, in a sense, the reverse. It is to start with the rules of probability theory alone and ask—relative to the six properties we consider—what is or is not possible. For this task, we need only EPR, Bell and Kochen–Specker. But we do mention later some other no-go results in QM (not needed to complete figure 1).

Two comments on the particular framework we present. First, we work with a single probability measure on a single space, where points in the space describe measurements on particles and outcomes of those measurements. An alternative—more conventional—approach would be to use a family of probability measures on a space describing outcomes only, with different probability measures corresponding to different measurements[4]. In fact, all our requirements are stated in terms of conditional probabilities: if such-and-such a measurement is made, then what is the probability of a certain outcome? The distinction between the approaches might therefore seem small—both involve families of probability measures. But it matters. If we had started from a family of probability measures rather than a single measure, we would not have been able to derive all the relationships between properties shown in figure 1 without making some additional assumptions[5]. So, formally, our approach is more parsimonious. Yet, it does add an ingredient at the conceptual level—namely, the existence of a probability measure prior to conditioning on measurements. This measure may be thought of as representing the perspective of a 'super-observer' who observes the experimenters as well as the outcomes of the experiments. Does the existence of such a measure contradict the free will of an experimenter in deciding what measurements to make? We do not think so—because, as we said, we work only with the conditionals. Still, even if it plays a very small role in our treatment, the idea of such a measure seems deserving of further study. We leave this as beyond the scope of the current paper.

A second choice we make in our framework is to treat only finite probability spaces. This involves a tradeoff. On one hand, finiteness allows us to avoid all measure-theoretic issues. On the other hand, as an assumption on the space in which a hidden variable lives, finiteness is undoubtedly restrictive. To be precise, the first of our two existence theorems needs only a finite space in any case, but, under finiteness, the second can treat only rational probabilities. (We sketch the extension of our second theorem, using an infinite space, to all probabilities.) Of course, finiteness makes our versions of the no-go theorems weaker.

We derive figure 1 in the body of this paper. Before that, though, let us offer a comment on its conceptual meaning in QM. The main message of the no-go theorems is that in building a hidden-variable theory, some properties that might be viewed as desirable—at least, *a priori*—have to be given up. But there is a choice of what to give up. Arguably, it is more a matter of metaphysics than physics as to what choice to make. The point of a formal treatment—as in figure 1—is to give a precise statement of what the options are. There is a basic three-way tradeoff. We can have:

[4] We are grateful to Shelly Goldstein and a referee who both pointed to this issue.

[5] A reader preferring the more conventional approach can simply make these additional assumptions. We give details in the following section.

(i) **Determinism.** (As we will explain, this comes in a strong or a weak form.) This says that randomness reflects only observer ignorance. Once hidden variables are introduced, there is no residual randomness in the universe.

(ii) **Parameter independence.** This says that when conducting an experiment on a system of particles, the outcome of a measurement on one of the particles does not depend on what measurements are performed on other particles. (The intuitive appeal of this property is that often the particles are widely separated.) This is a way of saying that the universe is local.

(iii) **λ-Independence.** This says that the nature of the particles—as determined by the value of a hidden variable—does not depend on the experiment conducted. There is, in this sense, no dependence between the observer and the observed.

Any one of these properties is consistent with the predictions of QM. So are certain combinations of properties, as figure 1 shows. But, however *a priori* reasonable they may seem, we cannot have all three properties—or even certain pairs of properties. There is an inherent tradeoff. This is an inescapable feature of QM.

The rest of this paper is organized as follows. Section 2 lays out the framework and basic definitions. Section 3 presents two existence theorems on hidden-variable models. Sections 4–6 present EPR [16, 1935], Bell [2, 1964] and Kochen–Specker [24, 1967] in our probability-theoretic framework. Section 7 mentions some other impossibility results in QM not covered in this paper.

2. The models and their properties

Here is the hidden-variable question in a bit more detail. Start with a model of an experiment done in QM. Sometimes, the experiment will consist of measurements performed on several entangled particles. Or, the experiment might involve several measurements performed on a single particle. The model describes the set-up and outcome of the experiment and so will be called an 'empirical model.' The question is whether one can find a hidden-variable model—i.e., a model involving additional 'hidden' variables—which is empirically equivalent to the first model and which has desired properties. By 'empirically equivalent' we mean that the two models make the same (probabilistic) prediction about outcomes.

Formally, we consider a space

$$\begin{aligned}\Psi = \{a, a', \ldots\} \times \{b, b', \ldots\} \times \{c, c', \ldots\} \cdots \\ \times \{A, A', \ldots\} \times \{B, B', \ldots\} \times \{C, C', \ldots\} \times \cdots .\end{aligned}$$

The variables $A, B, C, \ldots$ are measurements, and the variables $a, b, c, \ldots$ are associated outcomes of measurements. There might be several particles: Ann performs a measurement on her particle, Bob performs a measurements on his particle, ... Or, Ψ might describe a case where several measurements are performed on one particle. The definitions to come apply in either case. We take each of the spaces in Ψ to be finite, and suppose that Ψ is a finite product.

Let Λ be a finite space in which a hidden variable λ lives[6]. The overall space is then

$$\Omega = \Psi \times \Lambda.$$

Definition 2.1. *An empirical model is a pair* (Ψ, q)*, where* q *is a probability measure on* Ψ*. A hidden-variable model is a pair* (Ω, p)*, where* p *is a probability measure on* Ω*.*

[6] Throughout, we talk about one hidden variable. But we put no structure on the space in which the hidden variable lives. Of course, in the infinite case, a measurable structure would be needed.

Essentially, we employ the probability measures q and p once conditioned on one or more measurements. Still, as we said in the introduction, we cannot quite dispense with the unconditional q and p, and work instead with a family of probability measures indexed by measurements (one family for q and one for p). If we did, we would lose lemmas 2.1 and 2.4—and hence certain relationships in figure 1. These relationships seem intuitively correct to us, so we prefer a formalism in which they can be derived. With an indexed family of measures, we would have to impose rather than derive these relationships.

The models of definition 2.1 stand alone. However, we are also interested in stating when different models are equivalent:

Definition 2.2. *An empirical model* (Ψ, q) *and a hidden-variable model* (Ω, p) *are (empirically) equivalent if for all* $a, b, c, \ldots, A, B, C, \ldots,$

$$q(A, B, C, \ldots) > 0 \textit{ if and only if } p(A, B, C, \ldots) > 0,$$

and when both are non-zero,

$$q(a, b, c, \ldots | A, B, C, \ldots) = p(a, b, c, \ldots | A, B, C, \ldots).$$

Here, we write '$a, b, c, \ldots$' as a shorthand for the event

$$\{(a, b, c, \ldots)\} \times \{A, A', \ldots\} \times \{B, B', \ldots\} \times \{C, C', \ldots\} \times \cdots$$

in Ψ, or the event

$$\{(a, b, c, \ldots)\} \times \{A, A', \ldots\} \times \{B, B', \ldots\} \times \{C, C', \ldots\} \times \cdots \times \Lambda$$

in Ω, and similarly for other expressions. We will adopt this shorthand throughout.

The non-nullness condition is simply to ensure that any measurements $(A, B, C, \ldots)$ which are possible in the empirical model are also possible in the hidden-variable model under consideration, and vice versa. Without this condition, it would be hard to compare the two models.

We will often calculate $p(a, b, c, \ldots | A, B, C, \ldots)$, for $p(A, B, C, \ldots) > 0$, from the formula

$$p(a, b, c, \ldots | A, B, C, \ldots) = \sum_{\{\lambda : p(A, B, C, \ldots, \lambda) > 0\}} p(a, b, c, \ldots | A, B, C, \ldots, \lambda) p(\lambda | A, B, C, \ldots).$$

Substituting this in definition 2.2, we see that the idea of equivalence is to reproduce a given probability measure q on the space Ψ by averaging under a probability measure p on an augmented space Ω, where Ω includes a hidden variable. The measure p is then subject to various conditions (described below).

One more basic definition:

Definition 2.3. *Two hidden-variable models* (Ω, p) *and* $(\Omega, \overline{p})$ *are (empirically) equivalent if for all* $a, b, c, \ldots, A, B, C, \ldots,$

$$p(A, B, C, \ldots) > 0 \textit{ if and only if } \overline{p}(A, B, C, \ldots) > 0,$$

and when both are non-zero,

$$p(a, b, c, \ldots | A, B, C, \ldots) = \overline{p}(a, b, c, \ldots | A, B, C, \ldots).$$

Now, we move on to the different properties of hidden-variable models. Figure 2 repeats figure 1 (without the checks and crosses), as a preview of the properties and relationships we will consider.

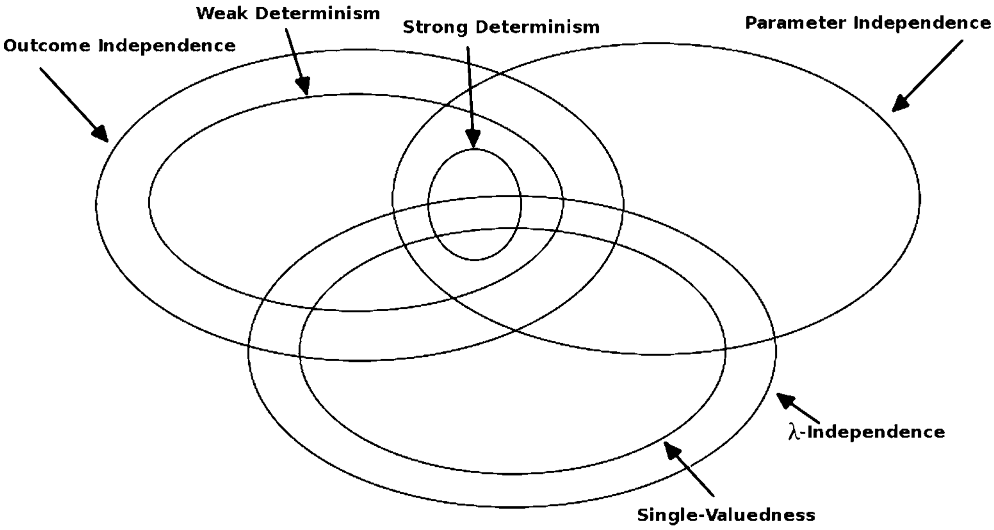

Figure 2. Properties of hidden-variable models.

Definition 2.4. *A hidden-variable model* (Ω, p) *satisfies single-valuedness if* Λ *is a singleton.*

This condition says that the hidden variable can take on only one value. In effect, this condition does not allow hidden variables. We include it because EPR will be usefully formulated this way.

Definition 2.5. *A hidden-variable model* (Ω, p) *satisfies* λ*-independence if for all* A, A', $B, B', C, C', \ldots, \lambda$, *whenever*

$$p(A, B, C, \ldots) > 0 \text{ and } p(A', B', C', \ldots) > 0,$$

then

$$p(\lambda|A, B, C, \ldots) = p(\lambda|A', B', C', \ldots).$$

(This term is from Dickson [14, 2005, p 140].) The condition says that the process determining the value of the hidden variable is independent of which measurements are chosen.

Remark 2.1. If a hidden-variable model satisfies single-valuedness, then it satisfies λ-independence.

Definition 2.6. *A hidden-variable model* (Ω, p) *satisfies strong determinism if, for every* A, λ, *whenever* $p(A, \lambda) > 0$, *there is an* a *such that* $p(a|A, \lambda) = 1$, *and similarly for* B, λ, b, *etc.*

Definition 2.7. *A hidden-variable model* (Ω, p) *satisfies weak determinism if, for every* $A, B, C, \ldots, \lambda$, *whenever* $p(A, B, C, \ldots, \lambda) > 0$, *there is a tuple* $a, b, c, \ldots$ *such that* $p(a, b, c, \ldots |A, B, C, \ldots, \lambda) = 1$.

Determinism is a basic condition in the literature. But we are careful to make a distinction between a strong and weak form. We will see that various results are true for one form but false for another. Broadly, the condition is that the hidden variable determines (almost surely) the outcomes of measurements. But, strong determinism says this holds measurement-by-measurement, while weak determinism says this holds only once all measurements are specified. There is a one-way implication:

Lemma 2.1. *If a hidden-variable model satisfies strong determinism, then it satisfies weak determinism.*

Proof. Suppose $p(A, B, C, \ldots, \lambda) > 0$. Then $p(A, \lambda) > 0$, $p(B, \lambda) > 0$, $p(C, \lambda) > 0, \ldots$. So, there are $a, b, c, \ldots$, such that $p(a|A, \lambda) = 1$, $p(b|B, \lambda) = 1$, $p(c|C, \lambda) = 1, \ldots$.

The result now follows from the following easy fact in probability theory: let $E_1, \ldots, E_n$ and $F_1, \ldots, F_n$ be events with $p\left(\bigcap_i F_i\right) > 0$. If $p(E_i|F_i) = 1$ for all i, then $p\left(\bigcap_i E_i \middle| \bigcap_i F_i\right) = 1$. □

Definition 2.8. *A hidden-variable model* (Ω, p) *satisfies outcome independence if for all* $a, b, c, \ldots, A, B, C, \ldots, \lambda$, *whenever* $p(A, B, C, \ldots, b, c, \ldots, \lambda) > 0$,

$$p(a|A, B, C, \ldots, b, c, \ldots, \lambda) = p(a|A, B, C, \ldots, \lambda), \tag{2.1}$$

and similarly with a and b interchanged, etc.

Outcome independence is taken from Jarrett [23, 1984] and Shimony [31, 1986]. It says that conditional on the value of the hidden variable and the measurements undertaken, the outcome of any one measurement is (probabilistically) unaffected by the outcomes of the other measurements.

Lemma 2.2. *A hidden-variable model* (Ω, p) *satisfies outcome independence if and only if for all* $a, b, c, \ldots, A, B, C, \ldots, \lambda$, *whenever* $p(A, B, C, \ldots, \lambda) > 0$,

$$\begin{aligned} p(a, b, c, \ldots |A, B, C, \ldots, \lambda) &= p(a|A, B, C, \ldots, \lambda) \times p(b|A, B, C, \ldots, \lambda) \\ &\quad \times p(c|A, B, C, \ldots, \lambda) \times \cdots. \end{aligned} \tag{2.2}$$

Proof. Standard; see, e.g., Chung [9, 1974, theorem 9.2.1]. □

Lemma 2.3 (Bub [7, 1997, p 69])**.** *If a hidden-variable model satisfies weak determinism, then it satisfies outcome independence.*

Proof. Suppose $p(A, B, C, \ldots, \lambda) > 0$. Then, by weak determinism, there is a tuple $a^*, b^*, c^*, \ldots$ such that

$$p(a, b, c, \ldots |A, B, C, \ldots, \lambda) = \chi_{\{a=a^*\}} \times \chi_{\{b=b^*\}} \times \chi_{\{c=c^*\}} \times \cdots.$$

But then $p(a|A, B, C, \ldots, \lambda) = \chi_{\{a=a^*\}}$, $p(b|A, B, C, \ldots, \lambda) = \chi_{\{b=b^*\}}$, $p(c|A, B, C, \ldots, \lambda) = \chi_{\{c=c^*\}}, \ldots$. Now use lemma 2.2. □

Definition 2.9. *A hidden-variable model* (Ω, p) *satisfies parameter independence if for all* $a, A, B, C, \ldots, \lambda$, *whenever* $p(A, B, C, \ldots, \lambda) > 0$,

$$p(a|A, B, C, \ldots, \lambda) = p(a|A, \lambda), \tag{2.3}$$

and similarly for $b, A, B, C, \ldots, \lambda$, *etc.*

Parameter independence is also from Jarrett [23, 1984] and Shimony [31, 1986]. It says that, conditional on the value of the hidden variable, the outcome of any one measurement depends (probabilistically) only on that measurement and not on the other measurements.

Lemma 2.4. *If a hidden-variable model satisfies strong determinism, then it satisfies parameter independence.*

Proof. Suppose $p(A, B, C, \ldots, \lambda) > 0$. Then $p(A, \lambda) > 0$. So, by strong determinism, there is an a^* such that $p(a|A, \lambda) = \chi_{\{a=a^*\}}$. But $p(a, A, B, C, \ldots, \lambda|A, \lambda) = p(a|A, B, C, \ldots, \lambda) \times p(A, B, C, \ldots, \lambda|A, \lambda)$, where $p(A, B, C, \ldots, \lambda|A, \lambda) > 0$. From $p(a^*|A, \lambda) = 1$, $p(a^*, A, B, C, \ldots, \lambda|A, \lambda) = p(A, B, C, \ldots, \lambda|A, \lambda)$. From $p(a|A, \lambda) = 0$ when $a \neq a^*$, we get $p(a, A, B, C, \ldots, \lambda|A, \lambda) = 0$. Thus, $p(a|A, B, C, \ldots, \lambda) = \chi_{\{a=a^*\}}$, establishing (2.3). □

Combinations of some of these properties give the well-known properties of locality and non-contextuality. First is locality, formulated by Bell [2, 1964]. In words, a hidden-variable model satisfies locality if the probability of getting some tuple of outcomes factorizes under the measurements.

Definition 2.10. *A hidden-variable model* (Ω, p) *satisfies locality if for all* $a, b, c, \ldots, A, B, C, \ldots, \lambda$*, whenever* $p(A, B, C, \ldots, \lambda) > 0$*,*

$$p(a, b, c, \ldots|A, B, C, \ldots, \lambda) = p(a|A, \lambda) \times p(b|B, \lambda) \times p(c|C, \lambda) \times \cdots. \tag{2.4}$$

Proposition 2.1 (Jarrett [23, 1984, p 582])**.** *A hidden-variable model satisfies locality if and only if it satisfies outcome independence and parameter independence.*

Proof. Assume $p(A, B, C, \ldots, \lambda) > 0$, and substitute (2.3) and its counterparts into (2.2). This yields (2.4).

Conversely, assume again $p(A, B, C, \ldots, \lambda) > 0$, and sum both sides of (2.4) over $b, c, \ldots$. This yields (2.3). Moreover, substituting (2.3) and its counterparts into (2.4) yields (2.2). □

Non-contextuality, due to Kochen–Specker [24, 1967], is a property of an empirical model. It says that the probability of obtaining a particular outcome of a measurement does not depend on the other measurements performed.

Definition 2.11. *An empirical model* (Ψ, q) *satisfies non-contextuality if for all* $a, A, B, B', C, C', \ldots$*, whenever* $q(A, B, C, \ldots) > 0$ *and* $q(A, B', C', \ldots) > 0$*,*

$$q(a|A, B, C, \ldots) = q(a|A, B', C', \ldots).$$

Also, the corresponding conditions must hold for $b, A, A', B, C, C', \ldots$*, etc.*

Proposition 2.2. *If a hidden-variable model* (Ω, p) *satisfies* λ*-independence and parameter independence, then any equivalent empirical model* (Ψ, q) *satisfies non-contextuality.*

Proof. We can assume $p(A, B, C, \ldots) > 0$ and $p(A, B', C', \ldots) > 0$. Then

$$\begin{aligned} p(a|A, B, C, \ldots) &= \sum_{\{\lambda: p(A,B,C,\ldots,\lambda)>0\}} p(a|A, B, C, \ldots, \lambda) p(\lambda|A, B, C, \ldots) \\ &= \sum_{\{\lambda: p(A,B,C,\ldots,\lambda)>0\}} p(a|A, B, C, \ldots, \lambda) p(\lambda) \\ &= \sum_{\{\lambda: p(A,B,C,\ldots,\lambda)>0\}} p(a|A, \lambda) p(\lambda), \end{aligned}$$

where the second line uses λ-independence and the third line uses parameter independence. Using $p(A, B, C, \ldots) > 0$ and λ-independence again, we have $p(A, B, C, \ldots, \lambda) > 0$ if and only if $p(\lambda) > 0$. So,

$$p(a|A, B, C, \ldots) = \sum_{\{\lambda: p(\lambda)>0\}} p(a|A, \lambda) p(\lambda).$$

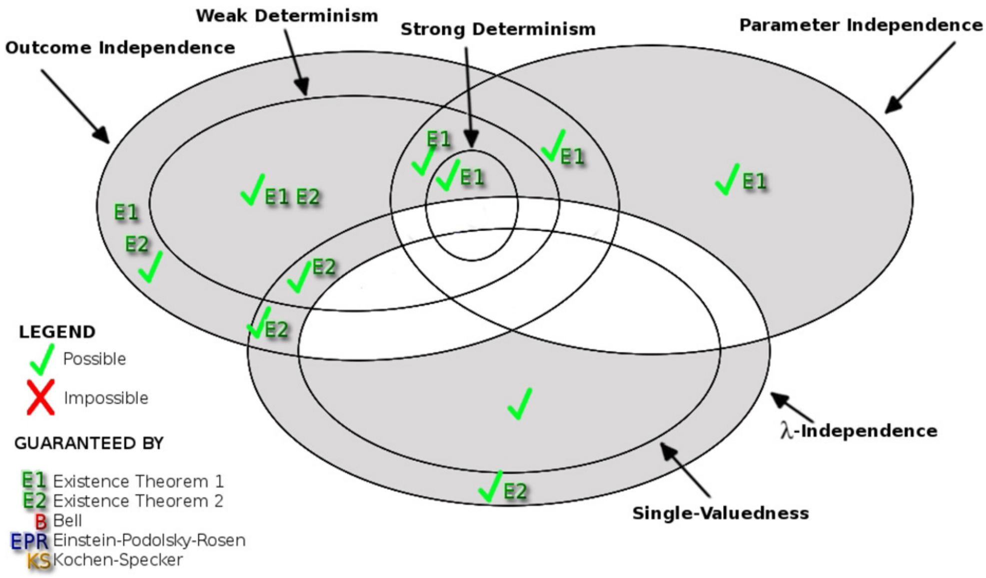

Figure 3. Existence theorems.

A similar argument establishes

$$p(a|A, B', C', \ldots) = \sum_{\{\lambda : p(\lambda)>0\}} p(a|A, \lambda)p(\lambda),$$

so that $p(a|A, B, C, \ldots) = p(a|A, B', C', \ldots)$, as required. □

3. Two existence theorems

We next prove two existence theorems for hidden-variable models which say what type of properties can always be found:

(E1) *Given any empirical model, there is an equivalent hidden-variable model which satisfies strong determinism.*

(E2) *Given any empirical model, there is an equivalent hidden-variable model which satisfies weak determinism and λ-independence.*

That is, each of these sets of conditions on a hidden-variable theory can always be satisfied. They cannot be impeded by any no-go theorems. Figure 3 repeats part of figure 1, putting a check in a region where there is always an equivalent hidden-variable model with the properties that hold in that region. The checks are followed by E1 and/or E2 which say which existence theorem pertains to that region.

(The region for single-valuedness alone also has a check. The existence of an equivalent hidden-variable model satisfying single-valuedness alone is immediate—it is essentially just the given empirical model. See remark 3.1 for a statement.)

Here are the two existence theorems. Similar methods to those in the first proof can be found in Fine [17, 1982, p 292]. The idea of the second theorem is in Teufel–Berndl–Dürr–Goldstein–Zanghì [33, 1997, p 1219] (see also Werner and Wolf [37, 2001, p 7]). But we have not found exact statements of the two theorems in the literature.

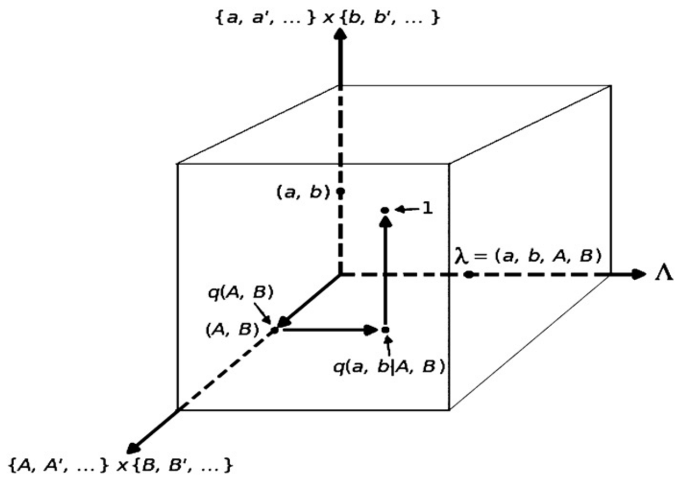

Figure 4. Construction for proof E1.

Theorem 3.1. *Given an empirical model* (Ψ, q)*, there is an equivalent hidden-variable model* (Ω, p) *which satisfies strong determinism.*

The proof is basically a mathematical trick. We simply take the hidden variable to be all the information possible. This means, in particular, that the hidden variable would have to 'know' the probabilities for different measurements and outcomes. With this huge hidden variable, we can build up the probability measure p from the given measure q. This construction is physically unsatisfying, of course—but not ruled out by the general concept of a hidden variable. It is also rather obvious. Bell [3, 1971] wrote: 'If no restrictions whatever are imposed on the hidden variables, or on the dispersion-free states, it is trivially clear that such schemes can be found to account for any experimental results whatever' (reprinted in [4, p 33]). Still, we give a proof—which, in particular, makes clear that even strong determinism is achieved. (This will not be possible in the next existence result.)

Proof. We give the proof for the case that Ψ is a 4-way product, but the extension to a general (finite) product will be clear. Set

$$\Lambda = \{a, a', \ldots\} \times \{b, b', \ldots\} \times \{A, A', \ldots\} \times \{B, B', \ldots\},$$

and define p in stages, as follows. (Figure 4 shows the construction.) For any pair A, B, set

$$p(A, B) = q(A, B). \tag{3.1}$$

For any pair A, B, and $\lambda = (\tilde{a}, \tilde{b}, \tilde{A}, \tilde{B})$, set

$$p(\lambda|A, B) = \begin{cases} q(\tilde{a}, \tilde{b}|A, B) & \text{if } \tilde{A} = A \text{ and } \tilde{B} = B, \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

For pairs a, b and A, B, and $\lambda = (\tilde{a}, \tilde{b}, \tilde{A}, \tilde{B})$, set

$$p(a, b|A, B, \lambda) = \begin{cases} 1 & \text{if } \tilde{a} = a, \tilde{b} = b, \tilde{A} = A, \tilde{B} = B, \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

This defines a measure p on Ω using

$$p(a, b, A, B, \lambda) = p(a, b|A, B, \lambda) \times p(\lambda|A, B) \times p(A, B).$$

(Note that $p(\cdot, \cdot|A, B, \lambda)$ is not a measure if $A \neq \tilde{A}$ or $B \neq \tilde{B}$. But then $p(\lambda|A, B) = 0$, so there is no difficulty.)

From (3.1), $p(A, B) > 0$ if and only if $q(A, B) > 0$. If both are positive, then from figure 4,

$$p(a, b|A, B) = 1 \times q(a, b|A, B),$$

so that equivalence is satisfied.

It remains to verify that (Ω, p) satisfies strong determinism. So, suppose $p(A, \lambda) > 0$. Writing $\lambda = (\tilde{a}, \tilde{b}, \tilde{A}, \tilde{B})$, we therefore assume $A = \tilde{A}$. Using (3.1)–(3.3),

$$p(a, A, \lambda) = p(a, \tilde{A}, \lambda) = \sum_{b', B'} p(a, b', \tilde{A}, B', \lambda) = p(a, \tilde{b}, \tilde{A}, \tilde{B}, \lambda) = q(a, \tilde{b}, \tilde{A}, \tilde{B}) \times \chi_{\{a=\tilde{a}\}},$$

$$p(A, \lambda) = p(\tilde{A}, \lambda) = \sum_{a', b', B'} p(a', b', \tilde{A}, B', \lambda) = p(\tilde{a}, \tilde{b}, \tilde{A}, \tilde{B}, \lambda) = q(\tilde{a}, \tilde{b}, \tilde{A}, \tilde{B}),$$

so that

$$p(a|A, \lambda) = \chi_{\{a=\tilde{a}\}},$$

which is strong determinism. □

Corollary 3.1. *Given a hidden-variable model (Ω, p), there is an equivalent hidden-variable model $(\Omega, \overline{p})$ which satisfies strong determinism.*

Proof. Start with (Ω, p), and (partially) define an equivalent empirical model (Ψ, q) by

$$q(a, b|A, B) = \sum_{\{\lambda: p(A,B,\lambda)>0\}} p(a, b|A, B, \lambda) p(\lambda|A, B),$$

for $p(A, B) > 0$. (There is no difficulty in completing the definition of q.)

By theorem 3.1, there is a hidden-variable model $(\Omega, \overline{p})$ which is equivalent to (Ψ, q) and which satisfies strong determinism. But $(\Omega, \overline{p})$ is also equivalent to (Ω, p). □

We state the next existence result for the case of rational probabilities, to enable us to prove it with a finite set Λ. After the proof, we sketch the extension to all probabilities.

Theorem 3.2. *Given an empirical model (Ψ, q) with rational probabilities, there is an equivalent hidden-variable model (Ω, p) which satisfies weak determinism and λ-independence.*

The idea of the proof is to have the hidden variable live on the (discretized) unit interval, and then to split up the interval according to relevant probabilities.

Proof. We give a proof for the case that Ψ is a 4-way product, but the argument clearly extends. Let p be arbitrary (but with full support) on $\{A, A', \ldots\} \times \{B, B', \ldots\}$. Fix a pair of settings (A_i, B_j) with $q(A_i, B_j) > 0$. For a pair of outcomes (a_k, b_l), write the conditional probability as

$$q(a_k, b_l|A_i, B_j) = \frac{r_{ijkl}}{s_{ijkl}}$$

for some integers $r_{ijkl} \geqslant 0$ and $s_{ijkl} > 0$.

Let Λ have N points, where

$$N = \prod_{i,j,k,l} s_{ijkl}.$$

We let p be uniform on N, and then form the product on $\{A, A', \ldots\} \times \{B, B', \ldots\} \times \Lambda$. Thus, λ-independence is satisfied.

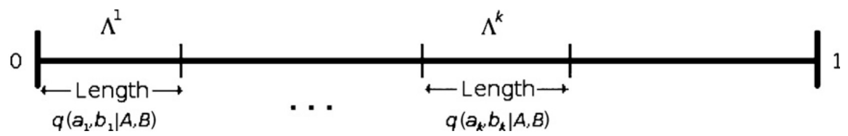

Figure 5. Construction for proof of extension of E2.

Still fixing (A_i, B_j), we 'assign' to (a_k, b_l) the number of points

$$r_{ijkl} \times \prod_{\substack{k' \neq k \\ l' \neq l}} s_{ijk'l'} \times \prod_{\substack{i' \neq i \\ j' \neq j}} \prod_{k,l} s_{i'j'kl}.$$

Formally, we mean that for each of these points,

$$p(a_k, b_l | A_i, B_j, \lambda) = 1.$$

Thus, weak determinism is satisfied. Note that, again for fixed (A_i, B_j), the probability of choosing one of these points is

$$\frac{r_{ijkl} \times \prod_{\substack{k' \neq k \\ l' \neq l}} s_{ijk'l'} \times \prod_{\substack{i' \neq i \\ j' \neq j}} \prod_{k,l} s_{i'j'kl}}{\prod_{i,j,k,l} s_{ijkl}} = \frac{r_{ijkl}}{s_{ijkl}},$$

and so

$$p(a_k, b_l | A_i, B_j) = \frac{r_{ijkl}}{s_{ijkl}} = q(a_k, b_l | A_i, B_j),$$

establishing equivalence. □

Corollary 3.2. *Given a hidden-variable model (Ω, p) with rational probabilities, there is an equivalent hidden-variable model $(\Omega, \overline{p})$ which satisfies weak determinism and λ-independence.*

Proof. As for corollary 3.1, using theorem 3.2 in place of theorem 3.1. □

We could handle irrational probabilities in theorem 3.2, if we allowed an infinite Λ. Set $\Lambda = [0, 1]$ and again take p to be uniform on Λ. Fix again a pair of settings (A, B) with $q(A, B) > 0$. Write the support of $q(\cdot, \cdot|A, B)$ as $\{(a_1, b_1), \ldots, (a_n, b_n)\}$. Partition Λ as in figure 5. On element Λ^k of the partition, we set $p(a_k, b_k|A, B, \lambda) = 1$. Conceptually, λ-independence and weak determinism follow as before. (But to make this formal, we would need to extend these definitions to infinite Λ.)

Finally in this section, we record the obvious fact:

Remark 3.1. Given an empirical model (Ψ, q), there is an equivalent hidden-variable model (Ω, p) which satisfies single-valuedness.

Proof. Let $\Lambda = \{\lambda\}$, and, for a, b, A, B, let $p(a, b, A, B, \lambda) = q(a, b, A, B)$. □

4. EPR

We have seen that certain types of hidden-variable model are always possible. Next come the no-go theorems, expressed in our framework. These show that no other type of hidden-variable model (among those covered by our six conditions) is necessarily possible.

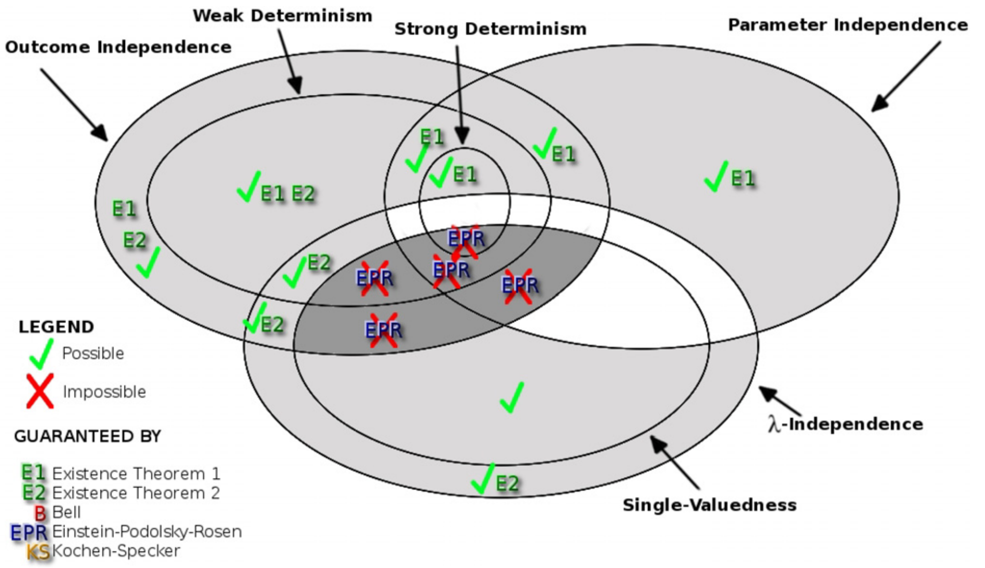

Figure 6. Existence theorems and EPR.

Here is the first no-go theorem, due to EPR [16, 1935], expressed in our framework. (Our formulation is very similar to that in Norsen [28, 2004].) Figure 6 adds crosses to figure 3, in accordance with the EPR result.

Note that, as with our other statements, EPR as given here is a simple result in probability theory. But the notation we use in the empirical model is meant to reflect the underlying physical set-up which was of interest to EPR. (More precisely, it reflects Bohm's [5, 1951] reformulation of EPR.) In the physical set-up, there are two entangled particles that are anti-correlated. If Ann measures positive spin, then Bob measures negative spin, and vice versa. There is a 50-50 chance of each pair of outcomes.

Theorem 4.1 (EPR [16, 1935])**.** *There is an empirical model* (Ψ, q) *for which there is no equivalent hidden-variable model* (Ω, p) *which satisfies single-valuedness and outcome independence.*

Proof. We let

$$\Psi = \{+_a, -_a\} \times \{+_b, -_b\} \times \{A\} \times \{B\},$$

and define q as in figure 7

	$+_b$	$-_b$
$+_a$	0	$\frac{1}{2}$
$-_a$	$\frac{1}{2}$	0

$q(\cdot, \cdot|A, B)$

Figure 7. Probabilities in EPR.

Now suppose, contra hypothesis, there is an equivalent hidden-variable model (Ω, p) satisfying single-valuedness and outcome independence. Let $\Lambda = \{\lambda\}$. Then we must have

$$p(+_a, -_b|A, B, \lambda) = p(-_a, +_b|A, B, \lambda) = \tfrac{1}{2},$$

from which

$$p(+_a|A, B, \lambda) = p(+_a, +_b|A, B, \lambda) + p(+_a, -_b|A, B, \lambda) = 0 + \tfrac{1}{2},$$

and

$$p(+_a|A, B, -_b, \lambda) = \frac{p(+_a, -_b|A, B, \lambda)}{p(-_b|A, B, \lambda)} = \frac{\frac{1}{2}}{\frac{1}{2}} = 1,$$

contradicting outcome independence. □

The conditions of EPR are tight. By remark 3.1, we cannot drop outcome independence. By theorem 3.1 or 3.2, we cannot drop single-valuedness. Here is a specific construction—for the EPR empirical model—of an equivalent hidden-variable model satisfying strong determinism (so, certainly outcome independence) and even λ-independence. Let $\Lambda = \{\lambda^1, \lambda^2\}$, and set $p(\lambda^1) = p(\lambda^2) = \frac{1}{2}$ and

$$p(+_a, -_b|A, B, \lambda^1) = 1, \qquad p(-_a, +_b|A, B, \lambda^2) = 1.$$

Using $p(A, B) = 1$, we see that the stated conditions hold.

At the level presented here, the EPR argument does not need any quantum effects. It could be realized entirely classically. Von Neumann [36, 1936] gave a nice example of classical action at a distance:

> Let S_1 and S_2 be two boxes. One knows that 1000 000 years ago either a white ball had been put into each or a black ball had been placed into each but one does not know which color the balls were. Subsequently one of the boxes (S_1) was buried on Earth, the other (S_2) on Sirius . . . . Now one digs S_1 on Earth out, opens it and sees: the ball is white. This action on Earth changes instantaneously the S_2 statistic on Sirius . . . .

In the QM context, EPR's conclusion was that the theory of QM needed to be 'completed.' This leads to the question of whether a construction like the one we just gave is always possible. This then leads to Bell's Theorem.

5. Bell

Bell's Theorem adds crosses to figure 6, as in figure 8.

Once more, our formulation is in probability terms alone. In the Bell experiment, Ann (respectively Bob) can make measurements of spin on her (respectively his) entangled particle in three directions. For each measurement, the only possible outcome is positive or negative spin. If the measurements are made in the same direction, the results will be anti-correlated (figure 9). Figure 10 gives the probabilities of the different outcomes of the measurements, when these are made in different directions. The probabilities in figure 10 are essentially quantum-mechanical.

Theorem 5.1 (Bell [2, 1964])**.** *There is an empirical model (Ψ, q) for which there is no equivalent hidden-variable model (Ω, p) which satisfies λ-independence, parameter independence and outcome Independence.*

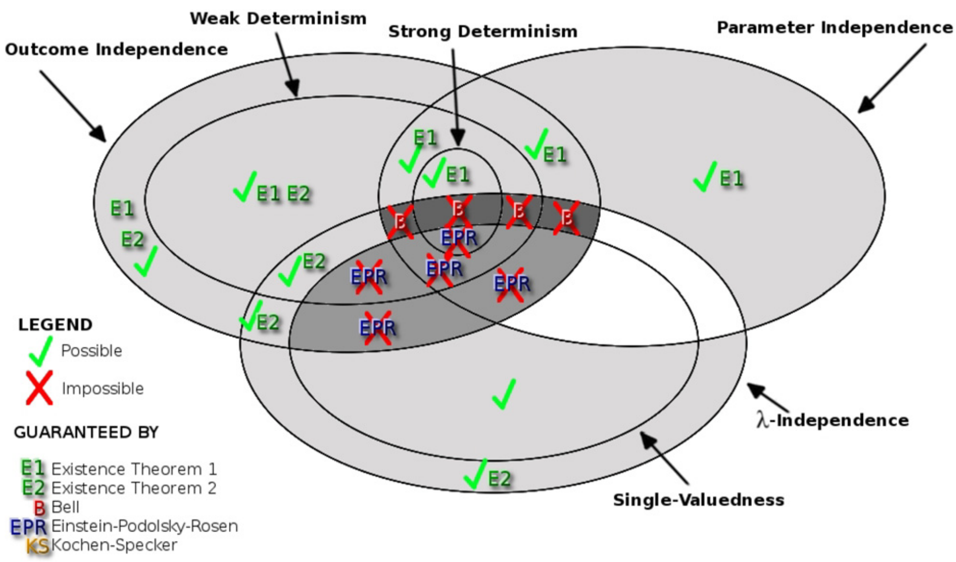

Figure 8. Existence theorems, EPR and Bell's Theorem.

	$+_b$	$-_b$
$+_a$	0	$\frac{1}{2}$
$-_a$	$\frac{1}{2}$	0

$q(\cdot,\cdot|A_i,B_i)$

Figure 9. Probabilities in Bell's Theorem.

Another phrasing (using proposition 2.1): there is no equivalent hidden-variable model which satisfies λ-independence and locality.

Proof. We let

$$\Psi = \{+_a, -_a\} \times \{+_b, -_b\} \times \{A_1, A_2, A_3\} \times \{B_1, B_2, B_3\},$$

and define q as in figures 9 and 10, with $q(A_i, B_j) = \frac{1}{9}$ for all i, j.

Now suppose, contra hypothesis, there is an equivalent hidden-variable model (Ω, p) satisfying λ-independence, parameter independence and outcome independence.

Fix an i. By assumption, $p(A_i, B_i) > 0$, since $q(A_i, B_i) > 0$. Using figure 9, we have

$$\begin{aligned} 0 = q(+_a, +_b|A_i, B_i) &= \sum_{\{\lambda : p(A_i, B_i, \lambda) > 0\}} p(+_a, +_b|A_i, B_i, \lambda) p(\lambda|A_i, B_i) \\ &= \sum_{\{\lambda : p(A_i, B_i, \lambda) > 0\}} p(+_a, +_b|A_i, B_i, \lambda) p(\lambda) \\ &= \sum_{\{\lambda : p(A_i, B_i, \lambda) > 0\}} p(+_a|A_i, \lambda) p(+_b|B_i, \lambda) p(\lambda), \end{aligned}$$

	$+_b$	$-_b$
$+_a$	$\frac{3}{8}$	$\frac{1}{8}$
$-_a$	$\frac{1}{8}$	$\frac{3}{8}$

$q(\cdot, \cdot|A_i, B_j)$ for $j \neq i$

Figure 10. Probabilities in Bell's Theorem (continued).

where the second line uses λ-independence and the third line uses parameter independence, outcome Independence and proposition 2.1. Using $p(A_i, B_i) > 0$ and λ-Independence again, we have $p(A_i, B_i, \lambda) > 0$ if and only if $p(\lambda) > 0$. Let $M = \{\lambda : p(\lambda) > 0\}$. Then,

$$p(+_a|A_i, \lambda) \times p(+_b|B_i, \lambda) = 0 \tag{5.1}$$

whenever $\lambda \in M$.

A similar argument using $q(-_a, -_b|A_i, B_i) = 0$ establishes

$$p(-_a|A_i, \lambda) \times p(-_b|B_i, \lambda) = 0 \tag{5.2}$$

whenever $\lambda \in M$.

Using (5.1) and (5.2), we see that for each i, there are disjoint sets $K_i, L_i \subseteq \Lambda$, with $K_i \cup L_i = M$, such that

$$\begin{array}{ll} p(+_a|A_i, \lambda) = 1 \text{ and } p(-_b|B_i, \lambda) = 1 & \text{when } \lambda \in K_i, \\ p(-_a|B_i, \lambda) = 1 \text{ and } p(+_b|B_i, \lambda) = 1 & \text{when } \lambda \in L_i. \end{array} \tag{5.3}$$

Similar to above, observe that

$$q(+_a, +_b|A_i, B_j) = \sum_M p(+_a|A_i, \lambda)p(+_b|B_j, \lambda)p(\lambda). \tag{5.4}$$

Using (5.3) (for i and j) in (5.4) we get

$$q(+_a, +_b|A_i, B_j) = p(K_i \cap L_j).$$

A parallel argument yields

$$q(-_a, -_b|A_i, B_j) = p(L_i \cap K_j).$$

Now use figure 10 to get

$$p(K_i \cap L_j) + p(L_i \cap K_j) = \tfrac{3}{4} \tag{5.5}$$

whenever $i \neq j$.

Refer to figure 11 (similar to figures in d'Espagnat [13, 1979]), and let

$$\begin{aligned} K_1 &= \boxed{1} \cup \boxed{4} \cup \boxed{5} \cup \boxed{8}, \\ L_1 &= \boxed{2} \cup \boxed{3} \cup \boxed{6} \cup \boxed{7}, \\ K_2 &= \boxed{1} \cup \boxed{2} \cup \boxed{5} \cup \boxed{6}, \\ L_2 &= \boxed{3} \cup \boxed{4} \cup \boxed{7} \cup \boxed{8}, \\ K_3 &= \boxed{1} \cup \boxed{2} \cup \boxed{3} \cup \boxed{4}, \\ L_3 &= \boxed{5} \cup \boxed{6} \cup \boxed{7} \cup \boxed{8}. \end{aligned}$$

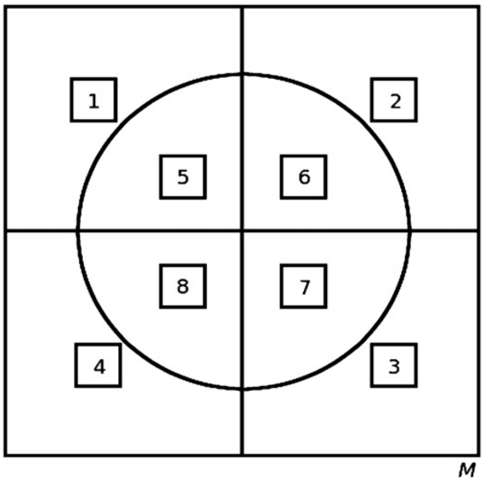

Figure 11. Construction for proof of Bell's Theorem.

Now (5.5) for $(i, j) = (1, 2), (2, 3)$ and $(3, 1)$ respectively, yields

$$p(K_1 \cap L_2) + p(L_1 \cap K_2) = \tfrac{3}{4},$$
$$p(K_2 \cap L_3) + p(L_2 \cap K_3) = \tfrac{3}{4},$$
$$p(K_3 \cap L_1) + p(L_3 \cap K_1) = \tfrac{3}{4},$$

or

$$p(\boxed{4}) + p(\boxed{8}) + p(\boxed{2}) + p(\boxed{6}) = \tfrac{3}{4}, \tag{5.6}$$

$$p(\boxed{5}) + p(\boxed{6}) + p(\boxed{3}) + p(\boxed{4}) = \tfrac{3}{4}, \tag{5.7}$$

$$p(\boxed{2}) + p(\boxed{3}) + p(\boxed{5}) + p(\boxed{8}) = \tfrac{3}{4}. \tag{5.8}$$

Adding (5.6)–(5.8) gives

$$2 \times (p(\boxed{2}) + p(\boxed{3}) + p(\boxed{4}) + p(\boxed{5}) + p(\boxed{6}) + p(\boxed{8})) = \tfrac{9}{4},$$

or

$$p(\boxed{2}) + p(\boxed{3}) + p(\boxed{4}) + p(\boxed{5}) + p(\boxed{6} + p(\boxed{8}) = \tfrac{9}{8},$$

which is impossible. □

Can we drop any of the conditions of Bell's Theorem? By theorem 3.1, we cannot drop λ-independence. By theorem 3.2, we cannot drop parameter independence.

For the Bell empirical model, we also cannot drop outcome Independence. To see this, let $\Lambda = \{\lambda\}$. Then λ-independence is satisfied. Define p on $\Psi \times \{\lambda\}$ from q on Ψ, as in remark 3.1. We then have

$$p(+_a|A_i, B_i, \lambda) = q(+_a|A_i, B_i) = \tfrac{1}{2} = q(+_a|A_i, B_j) = p(+_a|A_i, B_j, \lambda),$$
$$p(+_b|A_i, B_i, \lambda) = q(+_b|A_i, B_i) = \tfrac{1}{2} = q(+_b|A_j, B_i) = p(+_b|A_j, B_i, \lambda),$$

so that parameter independence is satisfied. Of course, outcome independence fails, as it must. For example

$$p(+_a|-_b, A_i, B_i, \lambda) = 1 \neq \tfrac{1}{2} = p(+_a|A_i, B_i, \lambda).$$

By contrast, the Kochen–Specker theorem produces an impossibility even without outcome independence.

Table 1.

A	E_1	E_1	E_8	E_8	E_2	E_9	E_{16}	E_{16}	E_{17}
B	E_2	E_5	E_9	E_{11}	E_5	E_{11}	E_{17}	E_{18}	E_{18}
C	E_3	E_6	E_3	E_7	E_{13}	E_{14}	E_4	E_6	E_{13}
D	E_4	E_7	E_{10}	E_{12}	E_{14}	E_{15}	E_{10}	E_{12}	E_{15}

6. Kochen–Specker

The Kochen–Specker [24, 1967] no-go result adds crosses to figure 8, to give a complete picture as in figure 1.

At the physical level, the Kochen–Specker experiment differs from those in the past two sections in considering measurements on only one particle. There are many presentations of Kochen–Specker, of course. We follow Cabello, Estebaranz and Garcìa–Alcaine [8, 1996], a simple treatment which results in the 4×9 array of table 1 (also presented in Held [22, 2000]). For various tuples of four orthogonal directions in 4-space (from a total of 18 directions), we ask whether or not the particle has spin in each of these directions. In each case, the answer will be that we get three directions without spin and only one direction with spin.

To state Kochen–Specker in our probabilistic framework, we will need to adapt the concept of exchangeability from probability theory (de Finetti [11, 1937], [12, 1972]). To give our definition, we consider the special case where the spaces of possible measurements are all the same, as are the spaces of possible outcomes:

$$\{A, \ldots\} = \{B, \ldots\} = \cdots = \{X_1, X_2, \ldots, X_m\},$$
$$\{a, \ldots\} = \{b, \ldots\} = \cdots = \{x_1, x_2, \ldots, x_n\},$$

for integers m, n. We will consider a permutation map π:

$$(A, B, \ldots) \mapsto (\pi(A), \pi(B), \ldots),$$
$$(a, b, \ldots) \mapsto (\pi(a), \pi(b), \ldots).$$

Note that we use π twice (despite the different domains), because we want to consider the same permutation on the two sequences.

Definition 6.1. *An empirical model* (Ψ, q) *satisfies exchangeability if for any indices* $i_1, i_2, \ldots \in \{1, 2, \ldots, m\}$ *and* $j_1, j_2, \ldots \in \{1, 2, \ldots, n\}$,

$$q(A = X_{i_1}, B = X_{i_2}, \ldots) > 0 \textit{ if and only if } q(\pi(A) = X_{i_1}, \pi(B) = X_{i_2}, \ldots) > 0,$$

for any permutation π, *and when both are non-zero,*

$$q(a = x_{j_1}, b = x_{j_2}, \ldots | A = X_{i_1}, B = X_{i_2}, \ldots)$$
$$= q(\pi(a) = x_{j_1}, \pi(b) = x_{j_2}, \ldots | \pi(A) = X_{i_1}, \pi(B) = X_{i_2}, \ldots).$$

In words, the requirement is that if we swap any number of measurements, then, as long as we swap the outcomes in the same way, the overall probability is unchanged. Thus, let q be the probability that Ann gets the outcome x_{j_1} and Bob gets the outcome x_{j_2}, if Ann performs measurement X_{i_1} on her particle and Bob performs measurement X_{i_2} on his particle. Let q' be the probability that Ann gets the outcome x_{j_2} and Bob gets the outcome x_{j_1}, if Ann performs measurement X_{i_2} on her particle and Bob performs measurement X_{i_1} on his particle. Exchangeability says that $q' = q$. Likewise, for several measurements on a single particle. This is similar to exchangeability à la de Finetti, though with a conditioning component.

Exchangeability might come from physical arguments. For example, the Bell model (figures 9 and 10) satisfies exchangeability. (This reflects the underlying physical fact that only the angle between the two measurements matters.)

Theorem 6.1 (Kochen–Specker [24, 1967])**.** *There is an empirical model* (Ψ, q) *for which there is no equivalent hidden-variable model that satisfies* λ*-independence and parameter independence.*

Kochen–Specker demonstrated the existence of a QM model that fails non-contextuality: whether or not their particle has spin in a certain direction is dependent on which other directions are also measured. The property of spin for such a particle does not stand alone. As the proof makes clear, theorem 6.1 is really a corollary to their result.

Proof. Consider an empirical model where

$$\{A, \ldots\} = \{B, \ldots\} = \{C, \ldots\} = \{D, \ldots\} = \{E_1, E_2, \ldots, E_{18}\},$$
$$\{a, \ldots\} = \{b, \ldots\} = \{c, \ldots\} = \{d, \ldots\} = \{0, 1\}.$$

Exchangeability is assumed to hold, and q assigns positive probability to each of the following nine tuples of measurement settings in table 1.

Finally, for any column, the empirical model has the property that precisely one of the following holds:

$$q\big(1, 0, 0, 0\big|E_{i_1}, E_{i_2}, E_{i_3}, E_{i_4}\big) = 1, \tag{6.1}$$
$$q\big(0, 1, 0, 0\big|E_{i_1}, E_{i_2}, E_{i_3}, E_{i_4}\big) = 1, \tag{6.2}$$
$$q\big(0, 0, 1, 0\big|E_{i_1}, E_{i_2}, E_{i_3}, E_{i_4}\big) = 1, \tag{6.3}$$
$$q\big(0, 0, 0, 1\big|E_{i_1}, E_{i_2}, E_{i_3}, E_{i_4}\big) = 1. \tag{6.4}$$

Now suppose, contra hypothesis, that there is an equivalent hidden-variable model satisfying λ-independence and parameter independence. By proposition 2.2, the above empirical model then satisfies non-contextuality.

Next take, say, the first column. If

$$q(0, 1, 0, 0|E_1, E_2, E_3, E_4) = 1, \tag{6.5}$$

then certainly

$$q(b = 1|E_1, E_2, E_3, E_4) = 1.$$

Since $(E_2, E_5, E_{13}, E_{14})$ is non-null, so is $(E_5, E_2, E_{13}, E_{14})$, by exchangeability. Using non-contextuality, we therefore have

$$q(b = 1|E_5, E_2, E_{13}, E_{14}) = 1,$$

from which, by exchangeability again,

$$q(a = 1|E_2, E_5, E_{13}, E_{14}) = 1.$$

Now use (6.1)–(6.4) to get

$$q(1, 0, 0, 0|E_2, E_5, E_{13}, E_{14}) = 1, \tag{6.6}$$

which tells us about the fifth column.

We therefore get a coloring problem: we try to color precisely one entry in each column—corresponding to the measurement that yields a 1. For example, suppose we color the entry E_2 in the first column—corresponding to (6.5). Then (6.6) tells us that we must color the entry E_2 in the fifth column. However, this is impossible. Each E_i appears an even number of times in table 1, and there is an odd number of columns. Thus, the table cannot be colored. □

7. Other no-go theorems

There are many important papers on the no-go question not touched upon here. These include Fine [17, 1982], [18, 1982], Greenberger, Horne, and Zeilinger [20, 1989], Malley and Fine [25, 2005], Mermin [26, 1990], [27, 1993], Peres [29, 1990], [30, 1991], and Szabo and Fine [32, 2002]. Again, our purpose is not to survey the literature. Rather, it is to give a complete picture of figure 1 and all its 21 regions. As figure 1 shows, just the three basic no-go theorems are needed for the six properties that we present.

The absence of Gleason's theorem ([19, 1957]) from our paper is a consequence of our choice not to impose any structure on our spaces (refer back to footnote 6). In particular, we do not work in Hilbert space. Of course, Gleason's theorem immediately implies the existence of the Kochen–Specker QM model (which we used in our theorem 6.1).

The recent no-go theorem of Conway and Kochen [10, 2006] generalizes Kochen–Specker by relaxing parameter independence. Consider a two-particle system. The requirement is that, conditional on the value of the hidden variable, the outcome of any particular measurement Ann makes on her particle may depend (probabilistically) on the other measurements she makes but not on the measurements Bob makes on his particle. We could accommodate this result by adding a seventh property—a generalized parameter independence—to our six, but refrain from pursuing this extension here.

Finally, we note the connection to Bohmian mechanics (Bohm [6, 1952]). Dürr–Goldstein–Zanghì [15, 2004, p 993] explain: 'In Bohmian mechanics the result obtained at one place at any given time will in fact depend upon the choice of measurement simultaneously performed at the other place.' Indeed, theorem 3.2 says that provided one is prepared to give up parameter independence, one can reproduce any empirical model—under λ-independence and weak determinism. Theorem 3.1 says that if one is prepared to give up λ-independence, one can get even strong determinism. In a sense, then, these results 'predict' the possibility of Bohmian mechanics—though not its specific content, of course.

Acknowledgments

We are indebted to Shelly Goldstein for detailed and valuable comments on an earlier draft. Our thanks to Amanda Friedenberg, Gus Stuart and three anonymous referees for important input, and also to Michael Giangrasso, Peter Morgan, Alex Peysakhovich and Jean de Valpine. Financial support from the Stern School of Business is gratefully acknowledged.

References

[1] Belinfante F 1973 *A Survey of Hidden-Variable Theories* (Oxford: Pergamon)
[2] Bell J 1964 On the Einstein–Podolsky–Rosen paradox *Physics* **1** 195–200 (Reprinted in [4, 1987, pp 14–21])
[3] Bell J 1971 Introduction to the hidden-variable question *Foundations of Quantum Mechanics Proc. of the International School of Physics 'Enrico Fermi', course IL* (NewYork: Academic) 171–81 (Reprinted in [4, 1987, pp 29–39])

[4] Bell J 1987 *Speakable and Unspeakable in Quantum Mechanics* (Cambridge: Cambridge University Press)
[5] Bohm D 1951 *Quantum Theory* (Englewood Cliffs, NJ: Prentice-Hall)
[6] Bohm D 1952 A suggested interpretation of the quantum theory in terms of 'hidden' variables, I and II *Phys. Rev.* **85** 166–193
[7] Bub J 1997 *Interpreting the Quantum World* (Cambridge: Cambridge University Press)
[8] Cabello A, Estebaranz J and Garcìa-Alcaine G 1996 Bell–Kochen–Specker theorem: a proof with 18 vectors *Phys. Lett.* A **212** 183–7
[9] Chung K L 1974 *A Course in Probability Theory* 2nd edn (NewYork: Academic)
[10] Conway J and Kochen S 2006 The free will theorem arXiv:quant-ph/0604079
[11] de Finetti B 1937 La Prévision: Ses Lois Logiques, Ses Sources Subjectives *Ann. de l'Institut Henri Poincaré* **7** 1–68
de Finetti B 1964 Foresight: its logical laws, its subjective sources *Studies in Subjective Probability* ed H Kyburg and H Smokler (NewYork: Wiley) pp 53–118 (Engl. Transl.)
[12] de Finetti B 1972 *Probability, Induction and Statistics* (NewYork: Wiley)
[13] d'Espagnat B 1979 The quantum theory and reality *Sci. Am.* 158–81
[14] Dickson W M 2005 *Quantum Chance and Non-Locality: Probability and Non-Locality in the Interpretations of Quantum Mechanics* (Cambridge: Cambridge University Press)
[15] Dürr D, Goldstein S and Zanghì N 2004 Quantum equilibrium and the role of operators as observables in quantum theory *J. Stat. Phys.* **116** 959–1055
[16] Einstein A, Podolsky B and Rosen N 1935 Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47** 770–80
[17] Fine A 1982 Hidden variables, joint probability, and the bell inequalities *Phys. Rev. Lett.* **48** 291–95
[18] Fine A 1982 Joint distributions, quantum correlations, and commuting observables *J. Math. Phys.* **23** 1306–10
[19] Gleason A 1957 Measures on closed sub-spaces of Hilbert spaces *J. Math. Mech.* **6** 885–93
[20] Greenberger D, Horne M and Zeilinger A 1989 Going beyond Bell's theorem *Bell's Theorem, Quantum Theory and Conceptions of the Universe* ed M Kafatos (Dordrecht: Kluwer) 73–6
[21] Gudder S 1988 *Quantum Probability* (NewYork: Academic)
[22] Held C 2000 The Kochen–Specker Theorem http://plato.stanford.edu/entries/kochen-specker
[23] Jarrett J 1984 On the physical significance of the locality conditions in the Bell arguments *Noûs* **18** 569–89
[24] Kochen S and Specker E 1967 The problem of hidden variables in quantum mechanics *J. Math. Mech.* **17** 59–87
[25] Malley J and Fine A 2005 Noncommuting observables and local realism *Phys. Lett.* A **347** 51–5
[26] Mermin N D 1990 Simple unified form for the major no-hidden-variables theorems *Phys. Rev. Lett.* **65** 3373–76
[27] Mermin N D 1993 Hidden variables and the two theorems of John Bell *Rev. Mod. Phys.* **65** 803–15
[28] Norsen T 2004 EPR and Bell locality arXiv:quant-ph/0408105
[29] Peres A 1990 Incompatible results of quantum measurements *Phys. Lett.* A **151** 107–8
[30] Peres A 1991 Two simple proofs of the Kochen–Specker theorem *J. Phys. A: Math. Gen.* **24** L175–8
[31] Shimony A 1986 Events and processes in the quantum world *Quantum Concepts in Space and Time* ed R Penrose and C Isham (Oxford: Oxford University Press) 182–203
[32] Szabo L and Fine A 1997 A local hidden variable theory for the GHZ experiment *Phys. Lett.* A **295** 229–40
[33] Teufel S, Berndl K, Dürr D, Goldstein S and Zanghì N 1997 Locality and causality in hidden variables models of quantum theory *Phys. Rev.* A **56** 1217–27
[34] van Fraassen B 1991 *Quantum Mechanics: An Empiricist View* (Oxford: Clarendon)
[35] von Neumann J 1932 *Mathematische Grundlagen der Quantenmechanik* (Berlin: Springer)
von Neumann J 1955 *Mathematical Foundations of Quantum Mechanics* (Princeton, NJ: Princeton University Press) (Engl. Transl.)
[36] von Neumann J 2005 Letter to E. Schrödinger *John von Neumann: Selected Letters (History of Mathematics* vol 27*)* ed M Rédei (Providence, RI: American Mathematical Society/London Mathematical Society) pp 211–3
[37] Werner R and Wolf M 2001 Bell inequalities and entanglement *Quantum Information & Comput.* **1** 1–25