

# OPTIMAL NO-GO THEOREM ON HIDDEN-VARIABLE PREDICTIONS OF EFFECT EXPECTATIONS

ANDREAS BLASS AND YURI GUREVICH

ABSTRACT. No-go theorems prove that, under reasonable assumptions, classical hidden-variable theories cannot reproduce the predictions of quantum mechanics. Traditional no-go theorems proved that hidden-variable theories cannot predict correctly the values of observables. Recent expectation no-go theorems prove that hidden-variable theories cannot predict the expectations of observables. We prove the strongest expectation-focused no-go theorem to date. It is optimal in the sense that the natural weakenings of the assumptions and the natural strengthenings of the conclusion make the theorem fail. The literature on expectation no-go theorems strongly suggests that the expectation-focused approach is more general than the value-focused one. We establish that the expectation approach is not more general.

## 1. INTRODUCTION

Hidden-variable theories allege that a state of a quantum system, even if it is pure and thus contains as much information as quantum mechanics permits, actually describes an ensemble of systems with distinct values of some hidden variables. Once the values of these variables are specified, the system becomes determinate or at least more determinate than quantum mechanics says. Thus the randomness in quantum predictions results, entirely or partially, from the randomness involved in selecting a member of the ensemble. No-hidden-variable theorems, in short no-go theorems, assert that, under reasonable hypotheses, no hidden-variable interpretation can reproduce the predictions of quantum mechanics.

The history of such no-go theorems goes back to John von Neumann's 1932 "Mathematische Grundlagen der Quantenmechanik" [18]. In this highly influential book, von Neumann gave a no-go argument, though he did not formulate it as a theorem. Von Neumann attempted to prove a no-go theorem, but made an error. In 1935, Grete Hermann spotted the error in the argument [12], but — as Mermin says [14] — she seems to have been entirely ignored. The error went virtually unnoticed until John Bell's 1966 paper [2] where he pointed out the error and proved

the first genuine no-go theorem. (The 1966 paper was written earlier, though published later, than Bell's 1964 paper [1].)

Independently of Bell, Simon Kochen and Ernst Specker published in 1967 their no-go theorem [13]. Their paper attracted the attention of philosophers, possibly because the authors were logicians and wrote in a way that appealed to philosophers, who then introduced the notion of contextuality. Here a context is a set of compatible observables, i.e. mutually commuting observables. Quantum theory is *contextual* in the sense that the value of an observable  $O$  measured as a part of one context may differ from the value of  $O$  measured as a part of another context. Kochen and Specker showed that a noncontextual hidden-variable theory cannot match the predictions of quantum mechanics for Hilbert spaces of dimension  $\geq 3$ . And of course classical theory is manifestly noncontextual; in fact, classically the value of an observable is the same whether it is measured all by itself or in company of any other observables.

In his 1993 paper [14], Mermin gave a lucid overview of the work of Bell and of Kochen-Specker, with useful simplifications and historical information. The section on von Neumann's error is called "von Neumann's silly assumption." Lately the story of that error has been challenged. According to the revisionists, von Neumann was misunderstood, and his implicit claim was more modest than a full-blown no-go theorem. Jeffrey Bub argues that von Neumann proved that quantum probabilities cannot be recovered from a hidden-variable theory of deterministic states [4]. Dennis Dieks argues that von Neumann proved that hidden-variable schemes violate the principle that physical quantities are representable by operators in a Hilbert space [8]. In a recent update of his 1993 paper [15], Mermin writes that he and Rüdiger Schack are preparing a reply to the revisionists. We are not going to dig deeper into the controversy here as it is tangential to the main issue of this paper.

The no-go theorems of Bell and of Kochen-Specker establish that, under suitable assumptions, hidden-variable theories cannot reproduce the predictions of quantum mechanics concerning *the possible values* obtained by measuring observables. There is no need to consider the probabilities of possible results or the expectation values of measurements; the measured values alone provide a discrepancy between hidden-variable theories and quantum theory. The hypotheses that are used to prove these theorems concern the measurements of observables in quantum states.

This value-focused approach leaves open the possibility of a hidden-variable theory that successfully predicts the expectations of observables. An alternative, expectation-focused approach was pioneered in 2008 by Spekkens [17]. In the new approach, the discrepancy between hidden-variable theories and quantum mechanics appears in the predictions of *the expected values of measurements*. There is no need to consider the actual values obtained by measurements or the probability distributions over these values. The hypotheses that are used to deduce these results concern the measurement of effects, i.e. the elements of positive operator-valued measurements (POVMs). Effects are represented by Hermitian operators with spectrum on the real interval  $[0, 1]$ . They are regarded as representing yes-or-no questions, the probability of “yes” for effect  $E$  in state  $|\psi\rangle$  being  $\langle\psi|E|\psi\rangle$ .

Earlier, in 2005, Spekkens proposed an operational definition of contextuality intended to generalize the traditional notion [16]. In particular, the definition of operational contextuality applies to state preparations and all other experimental procedures, not just projective measurements. It is this contextuality-centered research that led Spekkens to his expectation no-go theorem.

Spekkens’s expectation no-go theorem is restricted to finite-dimensional systems. In 2008–2009, Joseph Emerson and Christopher Ferrie strengthened Spekkens’s theorem [9, 10] and promoted frame representations of quantum mechanics. Their expectation no-go theorems are also restricted to finite-dimensional systems. In 2010, the third paper of Emerson and Ferrie on the subject, coauthored by Ryan Morris, generalized their results to infinite dimensional systems by means of a novel and informative proof [11].

The literature on the expectation approach [17, 9, 10, 11] makes the following two claims rooted in Spekkens’s generalization of contextuality which applies not only to measurements but also to state preparations.

**Generality:** The expectation approach is more general than the value approach.

**Symmetry:** The expectation approach is symmetric: it is even-handed in its treatment of state preparations and measurements.

The fact that operational contextuality applies to state preparations and measurements does not by itself imply the symmetry claim, but it invites it. And indeed all no-go theorems in the expectation-approach literature are even-handed in their treatment of state preparations and measurements.

The bold philosophical character of Spekkens's papers attracted our attention. We started with repairing mathematical flaws in the literature on expectation no-go theorems [5, 6]. In the process, we looked more closely at the generality claim. The operational definition of contextuality seems to ignore the commutation relation on observables which plays such an important role in traditional contextuality. At the end, as we show below, neither claim survived logical scrutiny.

In §2 of this paper, we formalize the notion of a hidden-variable theory predicting the expectations of effects. Contrary to preceding papers, we require only that the theory predicts the expectations of rank-1 projections, rather than the expectations of arbitrary effects. This leads to an expectation no-go theorem that is stronger than the earlier theorems in the literature. In fact, our no-go theorem is optimal in the sense that the natural weakenings of the assumptions and the natural strengthenings of the conclusion make the theorem fail.

The no-go theorem is proved in §3. It was partially motivated by our skepticism of the generality claim. One problem in comparing the two approaches is that they use different sorts of measurements. In the value approach, Hermitian operators serve as observables, and measuring one of them produces a number in its spectrum. In the expectation approach, certain Hermitian operators serve as effects, and measuring one of them produces 0 or 1, even if the spectrum consists entirely of other points. The only Hermitian operators for which these two uses coincide are projectors. We address this problem by weakening the hypotheses of the Ferrie-Morris-Emerson no-go theorem [11] so that they apply only to projectors, in fact to rank-1 projectors, rather than to arbitrary effects.

Our no-go theorem is not even-handed in its treatment of state preparations and measurements. Breaking the symmetry allowed us to prove a stronger theorem. In §4, we show that the expectation approach is not intrinsically symmetric.

In §5, toward comparing the expectation and value approaches, we formalize the value approach and prove a value-focused no-go theorem.

Finally, in §6, we point out that, while the operational contextuality is more general than the traditional one in the sense that it applies to state preparations and not only measurements, it is not more general in the sense of being implied by the traditional contextuality (as in the concept of mammal being more general than that of dog). Then we compare the expectation and value approaches. We argue that, while the expectation-focused approach has its advantages (in particular it works for two-dimensional Hilbert spaces), it is not more general than the value-focused one.

## 2. EXPECTATION REPRESENTATIONS

In this section we formalize the notion of hidden-variable theories predicting the expectations of effects.

Let  $\mathcal{H}$  be a Hilbert space, the state space of a quantum system. In this section, operators are by default linear operators on  $\mathcal{H}$ , and quantum states are density operators on  $\mathcal{H}$ . Thus our quantum states are in general mixed. The reason is that mixed states have physical meaning, and so it is desirable to include them in the theory.

We need some auxiliary definitions. As usual, a *convex linear combination* of vectors  $\vec{v}_1, \dots, \vec{v}_n$  in any vector space over the field  $\mathbb{C}$  of complex numbers has the form  $c_1\vec{v}_1 + \dots + c_n\vec{v}_n$  where the scalars  $c_i$  are nonnegative real numbers and  $c_1 + \dots + c_n = 1$ . Given a collection  $\mathcal{O}$  of linear operators on  $\mathcal{H}$ , a vector space  $V$  over  $\mathbb{C}$ , and a function  $f : \mathcal{O} \rightarrow V$ , we say that  $f$  is *convex linear* if it satisfies the following condition. For any  $A_0, \dots, A_n \in \mathcal{O}$ , if  $A_0$  is a convex linear combination  $c_1A_1 + \dots + c_nA_n$  of  $A_1, \dots, A_n$  then  $f(A_0) = c_1f(A_1) + \dots + c_nf(A_n)$ . We say that  $f$  is *weakly convex linear* if it satisfies the same condition for commuting operators  $A_1, \dots, A_n$ .

A positive self-adjoint operator  $E$  is an *effect* if the difference  $I - E$  between the identity operator  $I$  and  $E$  is positive as well. Let  $\mathcal{E}$  be a collection of effects on  $\mathcal{H}$ , and  $\mathcal{D}$  a collection of density operators on  $\mathcal{H}$ . The following definition describes hidden-variable theories over  $\mathcal{H}$  predicting the expectations of effects  $E \in \mathcal{E}$  in states  $\rho \in \mathcal{D}$ .

**Definition 2.1.** An *expectation representation* for  $\mathcal{H}, \mathcal{E}, \mathcal{D}$ , is a triple  $(\Lambda, F, \mu)$  where

- $\Lambda$  is a measurable space,
- $F$  is a weakly convex linear map which assigns to each effect  $E \in \mathcal{E}$  a measurable function  $F_E : \Lambda \rightarrow [0, 1]$ ,
- $\mu$  is a convex linear map which assigns to each state  $\rho \in \mathcal{D}$  a probability measure  $\mu_\rho$  on  $\Lambda$

such that

$$\int_{\Lambda} F_E d\mu_\rho = \text{Tr}(E\rho)$$

for all effects  $E \in \mathcal{E}$  and all states  $\rho \in \mathcal{D}$ .

The intention behind Definition 2.1 is as follows.

- $\Lambda$  is the range of a hidden variable  $\lambda$ . (If there are several hidden variables then we can view the tuple of all of them as a single variable.)

- $F_E(\lambda)$  is the expectation of  $E$  when the hidden variable takes the value  $\lambda$ . (Even when  $\lambda$  is fixed, there is still room for randomness in the values of effects.)
- If  $\Lambda$  is finite then  $\mu_\rho(\lambda)$  is the probability that the hidden variable takes value  $\lambda$  when the state is  $\rho$ ; accordingly  $\sum_\lambda F_E(\lambda)\mu_\rho(\lambda)$  is the overall expectation of  $E$  at  $\rho$ . In general,  $\mu_\rho$  is a probability measure on  $\Lambda$ , and  $\int_\Lambda F_E d\mu_\rho$  is the overall expectation of  $E$  at  $\rho$ .
- The equation  $\int_\Lambda F_E d\mu_\rho = \text{Tr}(E\rho)$  means that the overall expectation from the hidden variable theory matches the prediction of quantum theory.

A natural question arises why  $\mu$  is required to be convex linear in  $\rho$ . If you are a hidden-variable theorist, it is most natural for you to think of a mixed state as a classical probabilistic combination of component pure states. This leads you to the convex linearity of  $\mu$ . For example, if  $\rho = \sum_{i=1}^k p_i \rho_i$  where  $p_i$ 's are nonnegative reals and  $\sum p_i = 1$  then, by the rules of probability theory,  $\mu_\rho(S) = \sum p_i \mu_{\rho_i}(S)$  for any measurable  $S \subseteq \Lambda$ . Note, however, that you cannot start with any wild probability distribution  $\mu$  on pure states and then extend it to mixed states by convex linearity. There is an important constraint on  $\mu$  even on pure states. The same mixed state  $\rho$  may have different representations as a convex combination of pure states; all such representations must lead to the same probability measure  $\mu_\rho$ .

Motivated by Spekkens's idea of the even-handed treatment of preparations and measurements, the earlier literature on expectation no-go theorems [17, 9, 10, 11] required  $F$  to be convex linear on all effects. If an effect  $E_0$  is a linear combination  $c_1 E_1 + \dots + c_n E_n$  of commuting operators  $E_1, \dots, E_n$ , all these operators  $E_0, \dots, E_n$  can be measured simultaneously which justifies  $F(E_0) = c_1 F(E_1) + \dots + c_n F(E_n)$ . If these operators don't commute, then this equality does not seem physically justified to us. That is why we require only weak convex linearity of  $F$ .

### 3. EXPECTATION NO-GO THEOREM

In this section, we prove our optimal expectation-focused no-go theorem.

For any particular effect  $E$  on any Hilbert space  $\mathcal{H}$ , there is an expectation representation for  $\mathcal{H}$ , the singleton set  $\{E\}$  and all density operators on  $\mathcal{H}$ . Indeed, let  $\Lambda = \{0, 1\}$ ,  $F_E(\lambda) = \lambda$ ,  $\mu_\rho(1) = \text{Tr}(E\rho)$ , and  $\mu_\rho(0) = 1 - \text{Tr}(E\rho)$ ; then  $\int_\Lambda F_E d\mu_\rho = \text{Tr}(E\rho)$  for the fixed  $E$  and

all states  $\rho$ . But the following theorem shows that we cannot do the same for all effects simultaneously, not even for rank-1 projectors.

**Theorem 1** (Expectation no-go theorem). *If the dimension of the Hilbert space  $\mathcal{H}$  is at least 2 then there is no expectation representation for  $\mathcal{H}$ , all rank-1 projectors  $E$  on  $\mathcal{H}$ , and all density operators  $\rho$  on  $\mathcal{H}$ .*

The earlier expectation-focused no-go theorems [17, 9, 10, 11] assumed that  $F$  is convex linear and defined on all effects. We sharpen these theorems by assuming only that  $F$  is defined on rank-1 projectors and therefore vacuously convex linear. A rank-1 projector is never a convex linear combination of other rank-1 projectors.

Of course, we cannot expect any sort of no-go result in lower dimensions, because quantum theory in Hilbert spaces of dimensions 0 and 1 is trivial and therefore classical. Theorem 1 fails if the condition of convex linearity of  $\mu$  in Definition 2.1 is omitted; see Section 4.

The rest of this section is devoted to the proof of Theorem 1. We will assume the existence of an expectation representation and derive various consequences which will lead us to a contradiction.

We mentioned above that the third paper [11] of Ferrie and Emerson, coauthored by Ryan Morris, generalized their results to infinite dimensional systems. We achieve, in the following proposition, a similar generalization in a simpler way.

**Proposition 3.1** (Bootstrapping). *Let  $\mathcal{H}$  be a closed subspace of a Hilbert space  $\mathcal{H}'$ . Any expectation representation for  $\mathcal{H}'$ , all rank-1 projectors on  $\mathcal{H}'$  and all density operators on  $\mathcal{H}'$  gives rise to an expectation representation for  $\mathcal{H}$ , all rank-1 projectors on  $\mathcal{H}$  and all density operators on  $\mathcal{H}$ .*

*Proof.* We construct the desired expectation representation  $(\Lambda, F, \mu)$  for  $\mathcal{H}$  from a given representation  $(\Lambda', F', \mu')$  for the larger Hilbert space  $\mathcal{H}'$ . To begin, we set  $\Lambda = \Lambda'$ .

To define  $F$  and  $\mu$ , we use the inclusion map  $i : \mathcal{H} \rightarrow \mathcal{H}'$ , sending each element of  $\mathcal{H}$  to itself considered as an element of  $\mathcal{H}'$ , and we use its adjoint  $p : \mathcal{H}' \rightarrow \mathcal{H}$ , which is the orthogonal projector of  $\mathcal{H}'$  onto  $\mathcal{H}$ . Any state  $\rho$  over  $\mathcal{H}$ , gives rise to a state  $\bar{\rho} = i \circ \rho \circ p$  over  $\mathcal{H}'$ . Note that this expansion is very natural: If  $\rho$  corresponds to a pure state  $|\psi\rangle \in \mathcal{H}$ , i.e., if  $\rho = |\psi\rangle\langle\psi|$ , then  $\bar{\rho}$  corresponds to the same  $|\psi\rangle \in \mathcal{H}'$ . If, on the other hand,  $\rho$  is a mixture of states  $\rho_i$ , then  $\bar{\rho}$  is the mixture, with the same coefficients, of the  $\bar{\rho}_i$ . Define  $\mu_\rho = \mu'_{\bar{\rho}}$ .

Next we define  $F$ . For any rank-1 projector  $E$  in  $\mathcal{H}$ ,  $\bar{E} = i \circ E \circ p$  is a rank-1 projector in  $\mathcal{H}'$ , and so we define  $F_E = F'_{\bar{E}}$ . If  $E$  projects to

the one-dimensional subspace spanned by  $|\psi\rangle \in \mathcal{H}$ , then  $\bar{E}$  projects to the same subspace, now considered as a subspace of  $\mathcal{H}'$ .

This completes the definition of  $(\Lambda, F, \mu)$ . Most of the requirements in Definition 2.1 are trivial to verify. For the last requirement, the agreement between the expectation computed as a trace in quantum mechanics and the expectation computed as an integral in the expectation representation, it is useful to notice first that  $p \circ i$  is the identity operator on  $\mathcal{H}$ . We can then compute, for any state  $\rho$  and any rank-1 projector  $E$  on  $\mathcal{H}$ ,

$$\begin{aligned} \int_{\Lambda} F_E d\mu_{\rho} &= \int_{\Lambda} F'_{\bar{E}} d\mu'_{\bar{\rho}} = \text{Tr}(\bar{E}\bar{\rho}) = \text{Tr}(i \circ E \circ p \circ i \circ \rho \circ p) \\ &= \text{Tr}(i \circ E \circ \rho \circ p) = \text{Tr}(E \circ \rho \circ p \circ i) = \text{Tr}(E\rho), \end{aligned}$$

as required.  $\square$

By virtue of the Bootstrapping Proposition, we may assume without loss of generality that  $\mathcal{H}$  is finite dimensional. (We could in fact assume that  $\mathcal{H}$  is two-dimensional, but this more restrictive assumption would not make the rest of the proof any simpler.)

Suppose, toward a contradiction, that we have an expectation representation  $(\Lambda, F, \mu)$  for a finite-dimensional Hilbert space  $\mathcal{H}$  and the rank-1 projectors on  $\mathcal{H}$ . In the rest of this section, operators are by default linear operators on  $\mathcal{H}$ . Let  $\mathcal{S}$  be the vector space of self-adjoint operators.

**Lemma 3.1.** *The map  $\mu$  extends to a linear map on  $\mathcal{S}$ , which we also denote  $\mu$ , such that, for every  $A \in \mathcal{S}$ ,  $\mu_A$  is a bounded, signed, real-valued measure on  $\Lambda$  with the property that  $\int_{\Lambda} F_E d\mu_A = \text{Tr}(EA)$  for every rank-1 projector  $E$  on  $\mathcal{H}$ .*

*Proof.* Every self-adjoint operator  $A$  is the difference  $A_+ - A_-$  of two positive operators, where  $A_+$  has the same positive eigenvalues and corresponding eigenspaces as  $A$  but is identically zero on all the eigenspaces corresponding to non-positive eigenvalues.  $-A_-$  similarly matches the negative eigenvalues and eigenspaces of  $A$ ; we reverse its sign to get the positive operator  $A_-$ . It follows that positive operators span  $\mathcal{S}$ . Every positive operator of trace 1 is a state; as before we identify states with density operators. It follows that the states span  $\mathcal{S}$ . So it suffices to extend  $\mu$  to linear combinations of states.

If  $A = \sum_i a_i \rho_i$ , it is natural to extend  $\mu$  thus:  $\mu_A = \sum_i a_i \mu_{\rho_i}$ . If this definition is unambiguous, then we have

$$\int_{\Lambda} F_E d\mu_A = \sum_i a_i \int_{\Lambda} F_E d\mu_{\rho_i} = \sum_i a_i \text{Tr}(E\rho_i) = \text{Tr}(EA),$$



because the integral and the trace are linear.

To check that this extension is well defined, consider a self-adjoint operator  $A$  expressed in two distinct ways as a linear combination of states. Because we can pad our linear combinations with zero-coefficient summands, we may assume without loss of generality that our two linear combinations of states are linear combinations of the same states. So we have

$$\sum a_i \rho_i - \sum b_i \rho_i = A = \sum c_i \rho_i - \sum d_i \rho_i$$

for some coefficients  $a_i, b_i, c_i, d_i \geq 0$ . Therefore

$$\sum a_i \rho_i + \sum d_i \rho_i = \sum b_i \rho_i + \sum c_i \rho_i$$

Because all  $\rho_i$ 's have trace 1, it follows that  $\sum_i a_i + \sum_i d_i = \sum_i b_i + \sum_i c_i$ . Since the original linear combinations are distinct, not all coefficients  $a_i, b_i, c_i, d_i$  are zero; hence the number  $t = \sum_i a_i + \sum_i d_i = \sum_i b_i + \sum_i c_i > 0$ . We have

$$\sum \frac{a_i}{t} \rho_i + \sum \frac{d_i}{t} \rho_i = \sum \frac{b_i}{t} \rho_i + \sum \frac{c_i}{t} \rho_i,$$

where the left and right parts are positive self-adjoint operators of trace 1 and thus states. By the convex linearity of  $\mu$  on states,

$$\sum \frac{a_i}{t} \mu_{\rho_i} + \sum \frac{d_i}{t} \mu_{\rho_i} = \sum \frac{b_i}{t} \mu_{\rho_i} + \sum \frac{c_i}{t} \mu_{\rho_i}$$

which gives the desired

$$\sum a_i \mu_{\rho_i} - \sum b_i \mu_{\rho_i} = \sum c_i \mu_{\rho_i} - \sum d_i \mu_{\rho_i}. \quad \square$$

Let  $\mathcal{F}$  be the vector space of bounded measurable real-valued functions  $f$  on  $\Lambda$ .

**Corollary 3.1.** *For every  $f \in \mathcal{F}$ , the map  $A \mapsto \int_{\Lambda} f d\mu_A$  is a linear functional on  $\mathcal{S}$ .*

*Proof.* Use the linearity of  $\mu$  and of the integral.  $\square$

**Lemma 3.2.** *[Matching] Every  $f \in \mathcal{F}$  is matched by some self-adjoint operator  $[f]$  in the sense that  $\text{Tr}([f]A) = \int_{\Lambda} f d\mu_A$  for all  $A \in \mathcal{S}$ , i.e., that the functionals  $A \mapsto \int_{\Lambda} f d\mu_A$  and  $A \mapsto \text{Tr}([f]A)$  on  $\mathcal{S}$  coincide. Furthermore, the operator  $[f]$  is determined uniquely.*

The matching lemma is the key idea of the proof of the theorem. After elaborating on the nature of the map  $f \mapsto [f]$ , we will show that the structure of  $\mathcal{F}$  is too rich to be matched by that of  $\mathcal{S}$ .

*Proof.* Consider, for any  $B \in \mathcal{S}$ , the linear functional  $\bar{B}$  on  $\mathcal{S}$  defined by  $\bar{B}(A) = \text{Tr}(BA)$ . It suffices to show that every linear functional on  $\mathcal{S}$  is  $\bar{B}$  for a unique  $B \in \mathcal{S}$ . Indeed, then  $\int_{\Lambda} f d\mu_A = \bar{B}(A)$  for a unique  $B$ , which then serves as  $[f]$ .

The transformation  $T : B \rightarrow \bar{B}$  is linear. Since  $\mathcal{S}$  and its dual space  $\mathcal{S}^*$  have the same dimension, it suffices to show that  $\text{Kernel}(T) = 0$ . Suppose that  $B \in \text{Kernel}(T)$ . Then  $\text{Tr}(BA) = 0$  for every  $A \in \mathcal{S}$ . In particular  $\text{Tr}(BB^\dagger) = 0$ . But then  $B = 0$ .  $\square$

**Lemma 3.3.** *If  $B$  is self-adjoint operator such that  $\text{Tr}(B\rho) = \int_{\Lambda} f d\mu_{\rho}$  for all states  $\rho$  then  $B = [f]$ .*

*Proof.* By the linearity of the integral and trace,  $\text{Tr}(BA) = \int_{\Lambda} f d\mu_A$  for all  $A \in \mathcal{S}$ . By the Matching Lemma,  $B = [f]$ .  $\square$

**Corollary 3.2.**

- (1)  $[1] = I$  where  $1$  is the constant function with value 1 and  $I$  is the identity operator.
- (2)  $[F_E] = E$  for every rank-1 projector  $E$ .

*Proof.*

- (1) For every state  $\rho$ , we have  $\text{Tr}(I\rho) = 1 = \int_{\Lambda} 1 d\mu_{\rho}$ .
- (2) By the definition of expectation representation,  $\text{Tr}(E\rho) = \int_{\Lambda} F_E d\mu_{\rho}$  for all states  $\rho$ .  $\square$

We consider the standard partial orders on  $\mathcal{F}$  and on  $\mathcal{S}$ . The functions in  $\mathcal{F}$  are ordered pointwise. The operators in  $\mathcal{S}$  are ordered as follows:  $X \leq Y$  if  $Y - X$  is positive.

**Lemma 3.4.**

- (1) The transformation  $f \mapsto [f]$  from  $\mathcal{F}$  to  $\mathcal{S}$  is linear.
- (2) The linear transformation  $f \mapsto [f]$  is monotone: if  $f \leq g$  pointwise in  $\mathcal{F}$  then  $[f] \leq [g]$ , i.e. the operator  $[g] - [f]$  is positive.

*Proof.*

- (1)  $\text{Tr}([af + bg] \cdot A) = \int_{\Lambda} (af + bg) d\mu_A = a \int_{\Lambda} f d\mu_A + b \int_{\Lambda} g d\mu_A = a \text{Tr}([f] \cdot A) + b \text{Tr}([g] \cdot A)$ .
- (2) It suffices to suppose that a function  $h \in \mathcal{F}$  is nonnegative and prove that the operator  $[h]$  is positive, i.e. that  $\langle \psi|[h]|\psi \rangle \geq 0$  for all  $|\psi \rangle \in \mathcal{H}$ . The desired inequality is trivial if  $|\psi \rangle = 0$ , so let  $|\psi \rangle \neq 0$ . Normalizing  $|\psi \rangle$ , we may assume that its length is 1. Then  $|\psi \rangle \langle \psi|$  is a state and therefore  $\mu_{|\psi \rangle \langle \psi|}$  is a probability measure. In view of the definition of  $[h]$  and the positivity of  $h$ ,

$$\langle \psi|[h]|\psi \rangle = \text{Tr}([h]|\psi \rangle \langle \psi|) = \int_{\Lambda} h d\mu_{|\psi \rangle \langle \psi|} \geq 0. \quad \square$$

Now we are ready to complete the proof of Theorem 1. We show that the structure of  $\mathcal{F}$  given by the pointwise partial order and addition is too rich to be matched by the corresponding structure of  $\mathcal{S}$ .

For  $f, g \in \mathcal{F}$ , let  $(f \wedge g)(\lambda) = \min\{f(\lambda), g(\lambda)\}$ , so that  $(f + g - (f \wedge g))(\lambda) = \max\{f(\lambda), g(\lambda)\}$ . We restrict attention to functions in  $\mathcal{F}$  with values in the real interval  $[0, 1]$ ; recall that every  $F_E$  is such a function. We have

$$(1) \quad f + g - (f \wedge g) \leq 1.$$

Given any two rank-1 projectors  $A, B$  of  $\mathcal{H}$ , let  $f = F_A$ ,  $g = F_B$  and  $H = [f \wedge g]$ . By equation (1) and claim 2 of Lemma 3.4,

$$(2) \quad H \leq A, \quad H \leq B, \quad \text{and} \quad A + B - H \leq I.$$

Fix orthonormal vectors  $|0\rangle$  and  $|1\rangle$ . As usual, let  $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$  and  $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$ . Use equation (2) with  $A = |0\rangle\langle 0|$  and  $B = |+\rangle\langle +|$ . (Our choice of vectors  $|0\rangle$  and  $|+\rangle$  is convenient for the following calculation but any two non-orthogonal unit vectors would also work.) Since  $H$  and  $A - H$  are positive and  $A|1\rangle = 0$ , we obtain

$$\begin{aligned} 0 &\leq \langle 1|H|1\rangle \\ 0 &\leq \langle 1|(A - H)|1\rangle = \langle 1|A|1\rangle - \langle 1|H|1\rangle = -\langle 1|H|1\rangle \end{aligned}$$

so that  $\langle 1|H|1\rangle = 0$  and therefore, since  $H$  is positive,  $H|1\rangle = 0$ . Similarly, since  $H$  and  $B - H$  are positive and  $B|-\rangle = 0$ , we obtain  $H|-\rangle = 0$ . Being linear,  $H$  is identically zero on the subspace of  $\mathcal{H}$  spanned by  $|1\rangle$  and  $|-\rangle$ . Hence  $H|0\rangle = 0$ . By (2),

$$0 \leq \langle 0|(I - A - B + H)|0\rangle = \langle 0|0\rangle - \langle 0|A|0\rangle - \langle 0|B|0\rangle = 1 - 1 - \frac{1}{\sqrt{2}} = \frac{-1}{\sqrt{2}}.$$

This contradiction completes the proof of the theorem.

#### 4. ON THE SYMMETRY OF STATE PREPARATIONS AND MEASUREMENTS

In view of the idea of symmetry or even-handedness suggested by Spekkens [17], one might ask whether there is a dual version of Theorem 1, that is, a version that requires convex-linearity for effects but looks only at pure states and therefore does not require any convex-linearity for states.

The answer is no. With such requirements, for any  $\mathcal{H}$  there is a trivial expectation representation for  $\mathcal{H}$ , all effects on  $\mathcal{H}$  and all pure states on  $\mathcal{H}$ , so there cannot be a no-go theorem.

The example can be concisely described as taking the quantum state itself as the “hidden” variable. In more detail, fix a Hilbert state  $\mathcal{H}$  and consider an expectation representation  $(\Lambda, F, \mu)$  where

- $\Lambda$  is the set  $\{|\psi\rangle\langle\psi| : \psi \in \mathcal{H}\}$  of all pure states on  $\mathcal{H}$ .
- $F$  assigns to each effect  $E$  on  $\mathcal{H}$  the measurable function  $F_E(|\psi\rangle\langle\psi|) = \langle\psi|E|\psi\rangle$ .
- $\mu$  assigns to each pure state  $\rho = |\psi\rangle\langle\psi|$  the probability measure on  $\Lambda$  concentrated at the point  $|\psi\rangle\langle\psi|$ .

$\mu$  is vacuously convex-linear, and  $\int_{\Lambda} F_E d\mu_{\rho} = \text{Tr}(E\rho)$ . Notice that the map  $F$  is convex linear (in fact, linear) as a function of  $E$ .

It may seem that the map  $\mu$  can be extended to mixed states  $\rho = \sum p_i |\psi_i\rangle\langle\psi_i|$  by convex linearity:  $\mu_{\rho} = \sum p_i \mu_{|\psi_i\rangle\langle\psi_i|}$ . But the same mixed state, unless it is pure, has different representations as a convex combination of pure states. As a result a single  $\rho$  might be assigned different measures  $\mu_{\rho}$  by convex linearity. Theorem 1 shows that such a convex linear extension of  $\mu$  is impossible even if we restrict attention to effects that are rank-1 projectors.

This example also shows that Theorem 1 becomes false if the convex linearity requirement in Definition 2.1 is dropped.

## 5. THE VALUE APPROACH

Before we compare the expectation and value approaches, we need to define precisely what is expected of a hidden-variable theory in order for it to predict the correct values for observables. In this section, we give the necessary definition, and we use it to state and prove a general value-focused no-go theorem, based on the result of Kochen-Specker [13].

**Definition 5.1** (Value maps). *Let  $\mathcal{H}$  be a Hilbert space and  $\mathcal{O}$  a set of observables, i.e., self-adjoint operators on  $\mathcal{H}$ . A value map  $v$  for  $\mathcal{O}$  assigns to each  $A \in \mathcal{O}$  a number  $v(A)$  in the spectrum  $\sigma(A)$  of  $A$  in such a way that  $(v(A_1), \dots, v(A_n))$  is in the joint spectrum  $\sigma(A_1, \dots, A_n)$  of  $(A_1, \dots, A_n)$  whenever  $A_1, \dots, A_n$  are pairwise commuting.*

If  $\mathcal{H}$  is finite-dimensional then pairwise commuting operators  $A_1, \dots, A_n$  are simultaneously diagonalizable,  $\mathcal{H}$  admits an orthonormal basis of common eigenvectors of all the  $A_i$ 's, and the joint spectrum  $\sigma(A_1, \dots, A_n)$  consists of the tuples  $\nu = (\nu_1, \dots, \nu_n) \in \mathbb{R}^n$  that occur as the eigenvalues for such common eigenvectors. In general, even for infinite-dimensional  $\mathcal{H}$ , a point  $\nu \in \mathbb{R}^n$  belongs to  $\sigma(A_1, \dots, A_n)$  if and

only if for every  $\varepsilon \geq 0$  there is a unit vector  $|\psi\rangle \in \mathcal{H}$  (an approximate simultaneous eigenvector) such that  $\|A_i|\psi\rangle - \nu_i|\psi\rangle\| < \varepsilon$  for all  $i = 1, \dots, n$ ; see [3, Section 6.5].

The part of Definition 5.1 about pairwise commuting operators says exactly that, if one measures the observables  $A_1, \dots, A_n$  simultaneously, which is possible because they commute, then the values one obtains should be among the possibilities permitted by quantum mechanics, namely the  $n$ -tuples in the joint spectrum of the operators.

On the other hand, for observables that do not commute, quantum mechanics does not allow them to be simultaneously exactly measured, does not describe possible simultaneous values, and thus does not impose restrictions on value maps.

Traditionally, in the value-focused approach, states are pure and represented by unit vectors. We follow the tradition in this section as it suits our purpose here. We define value-focused hidden-variable theories.

**Definition 5.2.** A *value representation*  $\mathfrak{T}$  for a Hilbert space  $\mathcal{H}$  and a set  $\mathcal{O}$  of observables on  $\mathcal{H}$  assigns to each state  $|\psi\rangle \in \mathcal{H}$  a non-empty set  $\mathfrak{T}_{|\psi\rangle}$  of value maps for  $\mathcal{O}$  in such a way that, for every  $A \in \mathcal{O}$ , the set  $\{v(A) : v \in \mathfrak{T}_{|\psi\rangle}\}$  is exactly the set of possible values of  $A$  in state  $|\psi\rangle$  allowed by quantum mechanics.

A value representation treats a quantum state  $|\psi\rangle$  as an ensemble  $\mathfrak{T}_\psi$  of individual systems  $v$ , each of which has definite values  $v(A)$  for observables  $A \in \mathcal{O}$ . In the case where  $\mathcal{H}$  is finite-dimensional, for each observable  $A$ , the  $\{v(A) : v \in \mathfrak{T}_{|\psi\rangle}\}$  comprises the eigenvalues  $\lambda$  of  $A$  such that the projection of  $|\psi\rangle$  onto the eigenspace of  $\lambda$  is nonzero.

**Proposition 5.1** (Kochen-Specker). *In a Hilbert space of dimension 3 there is a finite set  $\mathcal{O}$  of rank-1 projectors for which no value map exists.*

*Proof.* The constructions given by Kochen and Specker [13] provide the desired  $\mathcal{O}$ . More precisely, the proof of their Theorem 1 uses a Boolean algebra generated by a finite set of one-dimensional subspaces of  $\mathcal{H}$ , and it shows that the projectors to those subspaces constitute an  $\mathcal{O}$  of the required sort.  $\square$

**Proposition 5.2.** *Let  $\mathcal{H}$  be a Hilbert space of dimension  $d \geq 3$ .*

- (1) *If  $d$  is infinite then there is a finite set  $\mathcal{O}$  of infinite-rank projectors of  $\mathcal{H}$  for which no value map exists.*
- (2) *If  $d$  is finite then there is a finite set  $\mathcal{O}$  of rank-1 projectors of  $\mathcal{H}$  for which no value map exists.*

*Proof.* If two Hilbert spaces are isomorphic and if one of them has a finite set  $\mathcal{O}$  of projectors with no value map, then the other also has a similar set; just conjugate the projectors in  $\mathcal{O}$  by any isomorphism between the two spaces. Thus, the existence of the desired set  $\mathcal{O}$  depends only on the dimension of  $\mathcal{H}$ , not on the specific space.

To prove (1), let  $\mathcal{K}, \mathcal{L}$  be Hilbert spaces with  $\dim(\mathcal{K}) = 3$  and  $\dim(\mathcal{L}) = d$ . The tensor product  $\mathcal{K} \otimes \mathcal{L}$  has the same dimension as  $\mathcal{H}$  and so it can be identified with  $\mathcal{H}$ . By the previous proposition, there is a set  $\mathcal{O}_1$  of rank-1 projectors in  $\mathcal{K}$  with no value map for  $\mathcal{O}_1$ . Define

$$\mathcal{O} = \{P \otimes I_{\mathcal{L}} : P \in \mathcal{O}_1\},$$

where  $I_{\mathcal{L}}$  is the identity operator on  $\mathcal{L}$ . The set  $\mathcal{O}$  of infinite-rank projectors has the algebraic structure of  $\mathcal{O}_1$ , so that there is no value map for  $\mathcal{O}$ .

To prove (2), we use induction on  $d \geq 3$ . The previous proposition provides the basis of induction. To handle the induction step, consider finite-dimensional Hilbert spaces  $\mathcal{H} \subseteq \mathcal{H}'$  where  $\dim(\mathcal{H}') = \dim(\mathcal{H}) + 1$  and suppose that  $\mathcal{H}$  has a finite set of rank-1 projectors for which no value map exists. We construct a finite set of rank-1 projectors in  $\mathcal{H}'$  for which no value map exists.

Let  $|\psi\rangle$  be any unit vector in  $\mathcal{H}'$ , and observe that its orthogonal complement,  $|\psi\rangle^\perp$ , is a subspace of  $\mathcal{H}'$  of the same dimension as  $\mathcal{H}$ . By the assumption, this subspace  $|\psi\rangle^\perp$  has a finite set  $\mathcal{O}$  of rank-1 projectors for which no value map exists. Each element  $P \in \mathcal{O}$  can be regarded as a rank-1 projector of  $\mathcal{H}'$ : if  $P = |\varphi\rangle\langle\varphi|$  in  $|\psi\rangle^\perp$ , interpret the same formula  $|\varphi\rangle\langle\varphi|$  in  $\mathcal{H}'$ .

Let  $\mathcal{O}_1$  consist of all the projectors from  $\mathcal{O}$ , interpreted as projectors of  $\mathcal{H}'$ , together with one additional rank-1 projector, namely  $|\psi\rangle\langle\psi|$ . What can a value map  $v$  for  $\mathcal{O}_1$  look like? It must send  $|\psi\rangle\langle\psi|$  to one of its eigenvalues, 0 or 1.

Suppose first that  $v(|\psi\rangle\langle\psi|) = 0$ . Then, using the fact that  $|\psi\rangle\langle\psi|$  commutes with all the other elements of  $\mathcal{O}_1$ , we easily compute that what  $v$  does to those other elements amounts to a value map for  $\mathcal{O}$ . But  $\mathcal{O}$  was chosen so that it has no value map, and so we cannot have  $v(|\psi\rangle\langle\psi|) = 0$ . Therefore  $v(|\psi\rangle\langle\psi|) = 1$ . (It follows that  $v$  maps the projectors associated to all the other elements of  $\mathcal{O}'$  to zero, but we shall not need this fact.)

We have thus shown that any value map for the finite set  $\mathcal{O}_1$  must send  $|\psi\rangle\langle\psi|$  to 1. Repeat the argument for another unit vector  $|\psi'\rangle$  that is orthogonal to  $|\psi\rangle$ . There is a finite set  $\mathcal{O}_2$  of rank-1 projectors such that any value map for  $\mathcal{O}_2$  must send  $|\psi'\rangle\langle\psi'|$  to 1. No value map

can send both  $|\psi\rangle\langle\psi|$  and  $|\psi'\rangle\langle\psi'|$  to 1, because their joint spectrum consists of only  $(1, 0)$  and  $(0, 1)$ . Therefore, there can be no value map for the union  $\mathcal{O}_1 \cup \mathcal{O}_2$ , which thus serves as the  $\mathcal{O}'$  required by the theorem.  $\square$

The assumption in part (2) that  $d$  is finite cannot be omitted. If  $d$  is infinite, then the set of all finite-rank projectors admits a value map, namely the constant zero function. This works because the definition of “value map” imposes constraints on only finitely many observables at a time.

**Theorem 2.** *Let  $\mathcal{H}$  be a Hilbert space of dimension  $d \geq 3$ . There is no value representation for  $\mathcal{H}$  and the set of all projectors on  $\mathcal{H}$ . If  $d$  is finite then there is no value representation for  $\mathcal{H}$  and the set of rank-1 projectors on  $\mathcal{H}$ .*

## 6. COMPARING THE TWO APPROACHES

The operational contextuality of Spekkens is more general than the traditional contextuality of Kochen-Specker in the sense that it applies not only to measurements but also to state preparations. This does not mean that if a no-go theorem rules out operationally noncontextual hidden-variable theories then it rules out traditionally noncontextual hidden-variable theories. Strengthening the hypothesis weakens the theorem. In fact, the opposite implication would be true if Spekkens just augmented the traditional measurement noncontextuality with a state-preparation noncontextuality.

But Spekkens didn’t simply add state-preparation noncontextuality to the traditional measurement noncontextuality. The operational version of measurement contextuality differs substantially from the traditional version. The traditional contextuality deals with the traditional measurements of observables producing eigenvalues and is subject to the quantum-theoretic restriction on the values of observables: If one measures simultaneously a tuple of commuting observables in a given state, the tuple of observed values is in the joint spectrum of these observables. The operational contextuality deals with the measurements of effects where the observed values need not be in the spectrum and thus the spectrum restriction need not hold. As a result, the two versions of contextuality are incompatible: neither implies the other. <sup>1</sup>

---

<sup>1</sup>By the way, we find that title claim “Negativity and contextuality are equivalent notions of nonclassicality” of [17] unsubstantiated. Spekkens observes that the notions of “noncontextual ontological model” and “nonnegative quasi-probability representation” are equivalent. His definition of “ontological model” includes a nonnegativity requirement (because it uses positive functions to model preparation

In the rest of this section, we compare the expectation-focused and value-focused approaches. It suffices for our purposes to restrict attention to the case where the Hilbert space  $\mathcal{H}$  in question is finite-dimensional. We argue that the expectation-focused approach is not more general than the value-focused one. But certain links between the two approaches are undeniable.

Consider an expectation representation  $\mathfrak{T} = (\Lambda, F, \mu)$  for a Hilbert space  $\mathcal{H}$ , the projectors on  $\mathcal{H}$  such that every  $F_P(\lambda) \in \{0, 1\}$ , and all pure states on  $\mathcal{H}$ . According to §4, such restricted expectation representations exist.

$\mathfrak{T}$  gives rise to a probabilistic value representation of sorts. For any  $\lambda \in \Lambda$ , the map  $P \mapsto F_P(\lambda)$  assigns to a projector  $P$  an eigenvalue of  $P$ , just as a standard value representation  $v$  would assign to an observable  $A$  an eigenvalue  $v(A)$  of  $A$ . We do not, however, impose on the map  $P \mapsto F_P(\lambda)$  the joint-spectrum requirement for commuting projectors. Also, for any pure state  $\rho$ , projector  $P$  and eigenvalue  $a$  of  $P$ , the probabilities  $\mu_\rho\{\lambda : F_P(\lambda) = a\}$  are determined correctly by the expectation  $\text{Tr}(P\rho)$  because  $F_P(\lambda)$  has only the two possible values 0 and 1 and because  $\mu_\rho\{\lambda : F_P(\lambda) = 1\} = \int_\Lambda F_\lambda(P)d\mu_\rho = \text{Tr}(P\rho)$ . Moreover,  $F$  admits an extension to arbitrary observables  $A$  so that for any pure state  $\rho$ , observable  $A$  and eigenvalue  $a$  of  $A$ , the probabilities  $\mu_\rho\{\lambda : F_A(\lambda) = a\}$  are correctly determined as well.

Notice, however, that this construction crucially depends on the fact that the joint-spectrum requirement is not imposed. This requirement is key in Kochen-Specker's contextuality proof, and it is absent in the expectation-focused approach.

The claim that the expectation-focused approach is more general than the value-focused one seems false. Certainly, there is no obvious way to prove Theorem 2 by means of the machinery developed to prove Theorem 1. But disproving the claim is difficult. In fact, the

---

procedures and measurement procedures), and his definition of “quasi-probability representation” implicitly includes a noncontextuality requirement (because it assigns functions directly to density operators and to POVMs, not to measurement procedures and preparation procedures). What he shows, therefore, is that adding noncontextuality to certain assumptions that already include nonnegativity produces the same notion as adding nonnegativity to other assumptions that already include noncontextuality. That does not make nonnegativity and noncontextuality equivalent. As an analogy, consider the algebraic notion of “field”. It can be obtained by adding commutativity to the notion of “division ring”, and it can also be obtained by adding the existence of inverses (for all nonzero elements) to the notion of “integral domain”. But the fact that “commutative division ring” is equivalent to “integral domain with inverses for all nonzero elements” does not make commutativity equivalent to the existence of inverses.



exact meaning of the claim is not clear. Theorem 2 does follow from Theorem 1 because any theorem follows from any proposition.

In principle, there are some ways to attack the claim. For example, one may try to construct a counter-factual physical world where the expectation no-go theorem holds but the value no-go theorem fails. This seems to be a thankless task. The burden of proof should be on those making the claim in the first place.

There is, however, one special case, that of dimension 2. Theorem 1 assumes  $\dim(\mathcal{H}) \geq 2$  while Theorem 2 assumes  $\dim(\mathcal{H}) \geq 3$ . So what about dimension 2?

Bell has given a hidden-variable theory for a two-dimensional Hilbert space in [2, §2], and so did Kochen and Specker in [13, §6]. Bell's theory was simplified by Mermin [14]. Building on this, we construct a value representation for  $\mathbb{C}^2$ , all observables on  $\mathbb{C}^2$ , and all pure states in  $\mathbb{C}^2$ . As usual  $\mathbb{C}^2$  is the Hilbert space of two-component vectors over the field  $\mathbb{C}$  of complex numbers.

**Theorem 3.** *There exists a value representation for  $\mathbb{C}^2$  and all observables on  $\mathbb{C}^2$ .*

In contrast, all expectation no-go theorems in the literature as well as Theorem 1 apply to the case of dimension 2. This is an important advantage of the expectation approach.

*Proof of Theorem 3.* Linear operators on  $\mathbb{C}^2$  are given by  $2 \times 2$  matrices over  $\mathbb{C}$ . Let  $\vec{\sigma}$  be the 3-component ‘‘vector’’ whose entries are the Pauli matrices  $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ,  $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ ,  $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ . Any observable  $A$  on  $\mathbb{C}^2$  has the form  $a_0 I + (\vec{a} \cdot \vec{\sigma})$  where  $a_0 \in \mathbb{R}$  and  $\vec{a} \in \mathbb{R}^3$ , and the eigenvalues of  $A$  are  $a_0 \pm \|\vec{a}\|$ . Two observables  $a_0 I + \vec{a} \cdot \vec{\sigma}$  and  $b_0 I + \vec{b} \cdot \vec{\sigma}$  commute if and only if  $\vec{a} = 0$  or  $\vec{b} = 0$  or else  $\vec{a} \neq 0$ ,  $\vec{b} \neq 0$  and  $\vec{a}, \vec{b}$  are collinear, i.e.,  $c\vec{a} = \vec{b}$  for some real coefficient  $c \neq 0$ . If the coefficient  $c > 0$  then  $\vec{a}, \vec{b}$  are parallel, and if  $c < 0$  then they are antiparallel. A map  $v$  that assigns to each observable  $A$  an eigenvalue of  $A$  is a value map if and only if the following two conditions hold for any two observables  $A = a_0 + \vec{a} \cdot \vec{\sigma}$ ,  $B = b_0 + \vec{b} \cdot \vec{\sigma}$  with nonzero vectors  $\vec{a}, \vec{b}$ :

- (1) if  $\vec{a}, \vec{b}$  are parallel then either  $v(A) = a_0 + \|\vec{a}\|$ ,  $v(B) = b_0 + \|\vec{b}\|$  or  $v(A) = a_0 - \|\vec{a}\|$ ,  $v(B) = b_0 - \|\vec{b}\|$ ,
- (2) if  $\vec{a}, \vec{b}$  are antiparallel then either  $v(A) = a_0 + \|\vec{a}\|$ ,  $v(B) = b_0 - \|\vec{b}\|$  or  $v(A) = a_0 - \|\vec{a}\|$ ,  $v(B) = b_0 + \|\vec{b}\|$ .

Let  $\mathcal{S}$  be the unit sphere in  $\mathbb{R}^3$ . If  $\vec{n} = (n_x, n_y, n_z) \in \mathcal{S}$ , then the dot product  $\vec{n} \cdot \vec{\sigma}$  is a Hermitian operator with eigenvalues  $\pm 1$ . Further, every state vector (i.e. every unit vector) in  $\mathbb{C}^2$  is an eigenvector, for

eigenvalue  $+1$ , of  $\vec{n} \cdot \vec{\sigma}$  for a unique  $\vec{n} \in \mathcal{S}$ ; we use the notation  $|\vec{n}\rangle$  for this eigenvector. If  $\mathbb{C}^2$  represents the states of a spin- $\frac{1}{2}$  particle, then the operator  $\frac{1}{2}\vec{n} \cdot \vec{\sigma}$  represents the spin component in the direction  $\vec{n}$ , and so  $|\vec{n}\rangle$  represents the state in which the spin is definitely aligned in the direction  $\vec{n}$ .

The desired value representation  $\mathfrak{T}$  assigns to each state  $|\vec{n}\rangle$  a set  $\mathfrak{T}_{|\vec{n}\rangle}$  of value maps

$$(3) \quad v_{|\vec{n}\rangle}(A, \vec{m}) = \begin{cases} a_0 + \|\vec{a}\| & \text{if } (\vec{m} + \vec{n}) \cdot \vec{a} \geq 0 \\ a_0 - \|\vec{a}\| & \text{if } (\vec{m} + \vec{n}) \cdot \vec{a} < 0. \end{cases}$$

where  $\vec{m}$  ranges over the punctured sphere  $\mathcal{S} - \{-\vec{n}\}$  and  $A = a_0I + \vec{a} \cdot \vec{\sigma}$  ranges over the observables on  $\mathbb{C}^2$ .

We need only check that, for any observable  $A = a_0 + \vec{a} \cdot \vec{\sigma}$  and any state  $|\vec{n}\rangle$ , the set  $\{v_{|\vec{n}\rangle}(A, \vec{m}) : \vec{m} \in \mathcal{S} - \{-\vec{n}\}\}$  comprises exactly the eigenvalues of  $A$  obtainable by measuring  $A$  in  $|\vec{n}\rangle$ . This is obvious if  $\vec{a} = 0$  so that  $A$  has a unique eigenvalue. Suppose that  $\vec{a} \neq 0$ . By definition,  $|\vec{n}\rangle$  is an eigenvector for a  $\vec{a} \cdot \vec{\sigma}$  if and only if  $\vec{n}$  is collinear with  $\vec{a}$ . Three cases arise.

Case 1:  $\vec{a} = \|\vec{a}\| \cdot \vec{n}$ . By definition,  $|\vec{n}\rangle$  is an eigenvector of  $\vec{n} \cdot \vec{\sigma}$  for eigenvalue  $+1$ , so that

$$A|\vec{n}\rangle = a_0|\vec{n}\rangle + (\vec{a} \cdot \vec{\sigma})|\vec{n}\rangle = a_0|\vec{n}\rangle + \|\vec{a}\|(\vec{n} \cdot \vec{\sigma})|\vec{n}\rangle = (a_0 + \|\vec{a}\|)|\vec{n}\rangle.$$

Thus  $|\vec{n}\rangle$  is an eigenvector of  $A$  for eigenvalue  $a_0 + \|\vec{a}\|$ ; any measurement of  $A$  in  $|\vec{n}\rangle$  returns  $a_0 + \|\vec{a}\|$ . It suffices to show that  $(\vec{m} + \vec{n}) \cdot \vec{a} = \vec{m} \cdot \vec{a} + \|\vec{a}\|$  is nonnegative for all  $\vec{m}$  in the sphere  $\mathcal{S}$ , which is obvious.

Case 2:  $\vec{a} = -\|\vec{a}\| \cdot \vec{n}$ . Since

$$A|\vec{n}\rangle = a_0|\vec{n}\rangle + (\vec{a} \cdot \vec{\sigma})|\vec{n}\rangle = a_0|\vec{n}\rangle - \|\vec{a}\|(\vec{n} \cdot \vec{\sigma})|\vec{n}\rangle = (a_0 - \|\vec{a}\|)|\vec{n}\rangle,$$

$|\vec{n}\rangle$  is an eigenvector of  $A$  for eigenvalue  $a_0 - \|\vec{a}\|$ , and any measurement of  $A$  in  $|\vec{n}\rangle$  returns  $a_0 - \|\vec{a}\|$ . We need to show that  $(\vec{m} + \vec{n}) \cdot \vec{a} = \vec{m} \cdot \vec{a} - \|\vec{a}\|$  is negative for all  $\vec{m}$  in the punctured sphere  $\mathcal{S} - \{\vec{n}\}$ , which is obvious.

Case 3:  $\vec{a}$  is not collinear with  $\vec{n}$ . In this case,  $|\vec{n}\rangle$  is not an eigenvector of  $A$ , and therefore both eigenvalues can be obtained by measuring  $A$  in  $|\vec{n}\rangle$ . It suffices to show that  $(\vec{m} + \vec{n}) \cdot \vec{a}$  is positive for some  $\vec{m} \in \mathcal{S} - \{\vec{n}\}$  and negative for some other  $\vec{m} \in \mathcal{S} - \{\vec{n}\}$ . It is positive when  $\vec{m} = \vec{a}/\|\vec{a}\|$ , and it is negative when  $\vec{m} = -\vec{a}/\|\vec{a}\|$ .  $\square$

Bell and Mermin had in mind a somewhat different hidden-variable theory.

**Proposition 6.1.** *There exists a hidden variable theory for  $\mathbb{C}^2$  which correctly predicts the expectation of any given observable in any given*

*pure state, as well as the probability that measuring a given observable in a given pure state results in a given eigenvalue.*

For the reader's convenience, we present Mermin's argument, somewhat simplified.

*Proof.* We use notation introduced in the proof of Theorem 3. The hidden variable  $\vec{m}$  ranges over the unit sphere  $\mathcal{S}$  in  $\mathbb{R}^3$  endowed with the uniform probability distribution. Any state  $|\vec{n}\rangle$  is treated as an ensemble  $v_{|\vec{n}\rangle}$  in (3) except that now  $\vec{m}$  ranges over all of  $\mathcal{S}$ , and in particular  $v_{|\vec{n}\rangle}(A, -\vec{n}) = a_0 + \|\vec{a}\|$ .

Given an observable  $A$  and a state  $|\vec{n}\rangle$ , we show (i) the average  $\int_{\mathcal{S}} v_{|\vec{n}\rangle}(A, \vec{m}) d\vec{m}$  is equal to the expectation  $\langle \vec{n} | A | \vec{n} \rangle$  and (ii) for every eigenvalue  $\alpha$  of  $A$ , the probability that  $v_{|\vec{n}\rangle}(A, \vec{m}) = \alpha$  is the probability that measuring  $A$  in  $|\vec{n}\rangle$  gives  $\alpha$ .

Since  $A$  has only two eigenvalues, its expectation in a given state determines their probabilities in that state; so (ii) follows from (i).

To prove (i), we may assume that  $a_0 = 0$ , because a general  $a_0$  would just be added to both sides of the equation. Thanks to the rotational symmetry of the situation (where rotations are applied to  $\vec{a}$ ,  $\vec{n}$  and  $\vec{m}$ ), we may assume that the vector  $\vec{a}$  points in the  $z$ -direction. Finally, by scaling, we may assume that  $\vec{a} = (0, 0, 1)$ . So our task is to prove that the average over  $\vec{m}$  of the values assigned to  $\sigma_z$  is  $n_z$ .

We have  $v_{|\vec{n}\rangle}(\sigma_z, \vec{m}) = \pm 1$  where the sign is chosen to agree with that of  $m_z + n_z$ . This  $m_z + n_z$  is the  $z$ -coordinate of a random point on the unit sphere centered at  $\vec{n}$ . So the question reduces to determining what fraction of this sphere lies above the  $x$ - $y$  plane. This plane cuts this unit sphere horizontally at a level  $n_z$  below the sphere's center. We invoke the theorem of Archimedes that a plane, cutting a sphere, divides the sphere's surface area in the same ratio that it divides the diameter perpendicular to the plane. So, by this theorem, our plane divides the sphere's area in the ratio of  $1 + n_z$  (above the plane) to  $1 - n_z$  (below the plane). That is, the value assigned to  $\sigma_z$  is  $+1$  with probability  $(1 + n_z)/2$  and  $-1$  with probability  $(1 - n_z)/2$ . Thus, the average value of  $\sigma_z$  is  $n_z$ , as required.  $\square$

What happens if we try to fit the hidden-variable theory in the proof of Proposition 6.1 into the form  $(\Lambda, F, \mu)$  of an expectation representation? A natural choice for  $\Lambda$  is the ball  $\{\vec{m} + \vec{n} : \vec{m}, \vec{n} \in \mathcal{S}\}$  of radius 2 centered at the origin of  $\mathbb{R}^3$ .  $F$  assigns to every observable  $A = a_0 + \vec{a} \cdot \vec{\sigma}$  the measurable function  $F_A(\vec{m} + \vec{n})$  equal  $a_0 + \|\vec{a}\|$  if  $(\vec{m} + \vec{n}) \cdot \vec{a} \geq 0$  and equal  $a_0 - \|\vec{a}\|$  otherwise. Finally,  $\mu$  assigns to each pure state  $|\vec{n}\rangle \langle \vec{n}|$  the

uniform distribution on the two-dimensional surface of a unit sphere centered at  $\vec{n}$ , because we are choosing  $\vec{m}$  randomly while  $\vec{n}$  is fixed.

If we restrict attention to projectors (rather than arbitrary observables) then our  $(\Lambda, F, \mu)$  is a genuine expectation representation for  $\mathbb{C}^2$ , all projectors on  $\mathbb{C}^2$  and all pure states on  $\mathbb{C}^2$ . (The proof of this fact uses that no projector is a nontrivial convex combination of others.) It is tempting to extend the expectation representation to all density operators by convex linearity. By Theorem 1, no such extension exists. Here is an example showing what goes wrong.

Consider the four pure states corresponding to spin in the directions of the positive  $x$ , negative  $x$ , positive  $z$ , and negative  $z$  axes. The corresponding density operators are the projectors

$$\frac{I + \sigma_x}{2}, \quad \frac{I - \sigma_x}{2}, \quad \frac{I + \sigma_z}{2}, \quad \frac{I - \sigma_z}{2},$$

respectively. Averaging the first two with equal weights, we get  $\frac{1}{2}I$ ; averaging the last two gives the same result. Since  $\mu$  is supposed to be convex linear, any extension would have to assign to the density operator  $\frac{1}{2}I$  the average of the probability measures assigned to the pure states with spins in the  $\pm x$  directions and also the average of the probability measures assigned to pure states with spins in the  $\pm z$  directions. But these two averages are visibly very different. The first is uniform on the union of two unit spheres tangent to the  $y$ - $z$ -plane at the origin, while the second is uniform on the union of two unit spheres tangent to the  $x$ - $y$ -plane at the origin.

## REFERENCES

- [1] John S. Bell, “On the Einstein-Podolsky-Rosen paradox,” *Physics* 1 (1964) 195–200.
- [2] John S. Bell, “On the problem of hidden variables in quantum mechanics,” *Reviews of Modern Physics* 38 (1966) 447–452.
- [3] Michael S. Birman and Michael Z. Solomjak, “Spectral theory of self-adjoint operators in Hilbert space,” Reidel 1987 (originally in Russian, Leningrad University Press 1980).
- [4] Jeffrey Bub, “von Neumann’s ‘No hidden variables’ proof: A re-appraisal,” *Foundations of Physics* 40 (2010) 1333–1340.
- [5] Andreas Blass and Yuri Gurevich, “Spekkens’s symmetric no-go theorem,” arXiv:1503.08084.
- [6] Andreas Blass and Yuri Gurevich, “On hidden variables: Value and expectation no-go theorems,” arXiv:1509.06896.
- [7] Andreas Blass and Yuri Gurevich, “Common denominator for value and expectation no-go theorems,” arXiv:1707.07368.

- [8] Dennis Dieks, “von Neumann’s impossibility proof: Mathematics in the service of rhetorics,” *Studies in History and Philosophy of Science Part B* 60 (2017) 136–148. Also arXiv:1801.09305v1.
- [9] Christopher Ferrie and Joseph Emerson, “Frame representations of quantum mechanics and the necessity of negativity in quasi-probability representations,” *Journal of Physics A: Mathematical and Theoretical* 41 352001 (2008). Also arXiv:0711.2658.
- [10] Christopher Ferrie and Joseph Emerson, “Framed Hilbert space: hanging the quasi-probability pictures of quantum theory,” *New Journal of Physics* 11 063040 (2009). Also arXiv:0903.4843.
- [11] Christopher Ferrie, Ryan Morris and Joseph Emerson, “Necessity of negativity in quantum theory,” *Physical Review A* 82, 044103 (2010). Also arXiv:0910.3198.
- [12] Grete Hermann, “Die naturphilosophischen Grundlagen der Quantenmechanik,” *Abhandlungen der Friesschen Schule*, 6 (1935), 75-152. English translation: Chapter 15 of “Grete Hermann: Between Physics and Philosophy,” E. Crull and G. Bacciagaluppi (eds.), Springer, 2017.
- [13] Simon Kochen and Ernst Specker, “The problem of hidden variables in quantum mechanics,” *Journal of Mathematics and Mechanics* 17 (1967) 59–87.
- [14] N. David Mermin, “Hidden variables and the two theorems of John Bell,” *Reviews of Modern Physics* 65 (1993) 803–815.
- [15] An update of [14], arXiv:1802.10119.
- [16] Robert W. Spekkens, “Contextuality for preparations, transformations, and unsharp measurements,” *Physical Review A* 71 (2005) 052108.
- [17] Robert W. Spekkens, “Negativity and contextuality are equivalent notions of nonclassicality,” *Physical Review Letters* 101(2) (2008) 020401. Also arXiv:0710.5549.
- [18] John von Neumann, “*Mathematische Grundlagen der Quantenmechanik*,” Springer, 1932. English translation: “*Mathematical Foundations of Quantum Mechanics*,” Princeton University Press, 1955.

MATHEMATICS DEPARTMENT, UNIVERSITY OF MICHIGAN, ANN ARBOR, MI 48109–1043, U.S.A.

*E-mail address:* `ablass@umich.edu`

MICROSOFT RESEARCH, ONE MICROSOFT WAY, REDMOND, WA 98052, U.S.A.

*E-mail address:* `gurevich@microsoft.com`