

# Common denominator for value and expectation no-go theorems

Andreas Blass  
University of Michigan  
ablass@umich.edu

Yuri Gurevich  
Microsoft Research  
gurevich@microsoft.com

Hidden-variable (HV) theories allege that a quantum state describes an ensemble of systems distinguished by the values of hidden variables. No-go theorems assert that HV theories cannot match the predictions of quantum theory. The present work started with repairing flaws in the literature on no-go theorems asserting that HV theories cannot predict the expectation values of measurements. That literature gives one an impression that expectation no-go theorems subsume the time-honored no-go theorems asserting that HV theories cannot predict the possible values of measurements. But the two approaches speak about different kinds of measurement. This hinders comparing them to each other. Only projection measurements are common to both. Here, we sharpen the results of both approaches so that only projection measurements are used. This allows us to clarify the similarities and differences between the two approaches. Neither one dominates the other.

## 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-go theorems assert that, under reasonable assumptions, a hidden-variable interpretation cannot reproduce the predictions of quantum mechanics. In this paper, we examine two species of such theorems, *value* no-go theorems and *expectation* no-go theorems. The value approach originated in the work of Bell [1, 2] and of Kochen and Specker [9] in the 1960's. Value no-go theorems establish that, under suitable hypotheses, hidden-variable theories cannot reproduce the predictions of quantum mechanics concerning the possible results of the measurements of observables.

The expectation approach was developed in the last decade by Spekkens [12] and by Ferrie, Emerson, and Morris [6, 7, 8], with [8] giving the sharpest result. In this approach, the discrepancy between hidden-variable theories and quantum mechanics appears in the predictions of the expectation values of the measurements of effects, i.e. the elements of POVMs, positive operator-valued measures. There is no need to consider the actual values obtained by measurements or the probability distributions over these values.

In both cases, measurements are associated to Hermitian operators, but they are 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 projections.

We sharpen the results of both approaches so that only projection measurements are used. Regarding the expectation approach, we substantially weaken the hypotheses. We do not need arbitrary effects, but only rank-1 projections. Accordingly, we need convex-linearity only for the hidden-variable picture of states, not for that of effects. Regarding the value approach, it turns out that rank-1 projections are sufficient in the finite dimensional case but not in general. Finally, using a successful hidden-variable theory of John Bell for a single qubit, we demonstrate that the expectation approach does not subsume the value approach.

## 2 Expectation No-Go Theorem

**Definition 1.** An *expectation representation* for quantum systems described by a Hilbert space  $\mathcal{H}$  is a triple  $(\Lambda, \mu, F)$  where

- $\Lambda$  is a measurable space,
- $\mu$  is a convex-linear map assigning to each density operator  $\rho$  on  $\mathcal{H}$  a probability measure  $\mu(\rho)$  on  $\Lambda$ , and
- $F$  is a map assigning to each rank-1 projection  $E$  in  $\mathcal{H}$  a measurable function  $F(E)$  from  $\Lambda$  to the real interval  $[0, 1]$ .

It is required that for all density matrices  $\rho$  and all rank-1 projections  $E$

$$\text{Tr}(\rho \cdot E) = \int_{\Lambda} F(E) d\mu(\rho) \quad (1)$$

The convex linearity of  $\mu$  means that  $\mu(a_1\rho_1 + a_2\rho_2) = a_1\mu(\rho_1) + a_2\mu(\rho_2)$  whenever  $a_1, a_2$  are nonnegative real numbers with sum 1.

The definition of expectation representation is similar to Ferrie-Morris-Emerson's definition of the probability representation [8] except that (i) the domain of  $F$  contains only rank-1 projections, rather than arbitrary effects, and (ii) we do not (and cannot) require that  $F$  be convex-linear.

Intuitively an expectation representation  $(\Lambda, \mu, F)$  attempts to predict the expectation value of any rank-1 projection  $E$  in a given mixed state  $\rho$ . The hidden variables are combined into one variable ranging over  $\Lambda$ . Further,  $\mu(\rho)$  is the probability measure on  $\Lambda$  determined by  $\rho$ , and  $(F(E))(\lambda)$  is the probability of determining the effect  $E$  at the subensemble of  $\rho$  determined by  $\lambda$ . The left side of (1) is the expectation of  $E$  in state  $\rho$  predicted by quantum mechanics and the right side is the expectation of  $F(E)$  in the ensemble described by  $\mu(\rho)$ .

But why is  $\mu$  supposed to be convex linear? Well, mixed states have physical meaning and so it is desirable that  $\mu$  be defined on mixed states as well. 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 the component 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  $\mu$ . 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  $\mu(\rho)$ .

**Theorem 2** (First Bootstrapping Theorem). *Let  $\mathcal{H}$  be a closed subspace of a Hilbert space  $\mathcal{H}'$ . From any expectation representation for quantum systems described by  $\mathcal{H}'$ , one can directly construct such a representation for systems described by  $\mathcal{H}$ .*

*Proof.* We construct an expectation representation  $(\Lambda, \mu, F)$  for quantum systems described by  $\mathcal{H}$  from any such representation  $(\Lambda', \mu', F')$  for the larger Hilbert space  $\mathcal{H}'$ . To begin, we set  $\Lambda = \Lambda'$ .

To define  $\mu$  and  $F$ , 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 projection of  $\mathcal{H}'$  onto  $\mathcal{H}$ . Any density operator  $\rho$  over  $\mathcal{H}$ , gives rise to a density operator  $\bar{\rho} = i \circ \rho \circ p$  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  $\mu(\rho) = \mu'(\bar{\rho})$ .

The definition of  $F$  is similar. For any rank-1 projection  $E$  in  $\mathcal{H}$ ,  $\bar{E} = i \circ E \circ p$  is a rank-1 projection 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  $\Lambda$ ,  $\mu$ , and  $F$ . Most of the requirements in Definition 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 density operator  $\rho$  and any rank-1 projection  $E$  on  $\mathcal{H}$ ,

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

as required. □

**Theorem 3** (Expectation no-go theorem). *If the dimension of the Hilbert space  $\mathcal{H}$  is at least 2 then there is no expectation representation for quantum systems described by  $\mathcal{H}$ .*

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. By the First Bootstrapping Theorem, it suffices to prove Theorem 3 just in the case  $\text{Dim}(\mathcal{H}) = 2$ . But we find Ferrie-Morris-Emerson's proof that works directly for all dimensions [8] instructive, and we adjust it to prove Theorem 3. The adjustment involves adding some details and observing that a drastically reduced domain of  $F$  suffices. The adjustment also involves making a little correction. Ferrie et al. quoted an erroneous result of Bugajski [5] which needs some additional hypotheses to become correct. Fortunately for Ferrie et al., those hypotheses hold in their situation.

*Proof.* The proof involves several normed vector spaces.

- $\mathcal{B}$  is the real Banach space of bounded self-adjoint operators  $\mathcal{H} \rightarrow \mathcal{H}$  with norm  $\|A\| = \sup\{\|Ax\| : x \in A, \|x\| = 1\}$ .
- $\mathcal{F}$  is the real vector space of bounded, measurable, real-valued functions on  $\Lambda$  with norm  $\|f\| = \sup\{|f(\lambda)| : \lambda \in \Lambda\}$ .
- $\mathcal{M}$  is the real vector space of bounded, signed, real-valued measures on  $\Lambda$  with the total variation norm  $\|\mu\| = \mu_+(\Lambda) + \mu_-(\Lambda)$  where  $\mu = \mu_+ - \mu_-$  and  $\mu_+$  and  $\mu_-$  are positive measures with disjoint supports.
- $\mathcal{T}$  is the vector subspace of  $\mathcal{B}$  consisting of the trace-class operators. These are the operators  $A$  whose spectrum consists of real eigenvalues  $\alpha_i$  such that the sum  $\sum_i |\alpha_i|$  is finite; eigenvalues with multiplicity  $> 1$  are repeated in this list, and the continuous spectrum is empty or  $\{0\}$ . The sum

$\sum_i |\alpha_i|$  serves as the norm of  $A$  in  $\mathcal{T}$ . The sum  $\sum_i \alpha_i$  of eigenvalues themselves (rather than their absolute values) is the trace of  $A$ . Note that density operators are positive trace-class operators of trace 1.

In the rest of the proof, by default, operators, transformation and functionals are bounded and of course linear. Suppose, toward a contradiction, that we have an expectation representation  $(\Lambda, \mu, F)$  for some  $\mathcal{H}$  with  $\text{Dim}(\mathcal{H}) \geq 2$ .

**Lemma 4.**  *$\mu$  can be extended in a unique way to a transformation, also denoted  $\mu$ , from all of  $\mathcal{T}$  into  $\mathcal{M}$ .*

*Proof of Lemma 4.* Every  $A \in \mathcal{T}$  can be written as a linear combination of two density operators. Indeed, if  $\|A\| > 0$  and  $A$  is positive then  $\text{Tr}(A) = \|A\|$  and  $A = \|A\|\rho$  where  $\rho = \frac{A}{\|A\|}$ . In general, it suffices to represent  $A$  as the difference  $B - C$  of positive trace-class operators. Choose  $A_+$  (resp.  $A_-$ ) to have the same positive (resp. negative) eigenvalues and corresponding eigenspaces as  $A$  and be identically zero on all the eigenspaces corresponding to the remaining eigenvalues. The desired  $B = A_+$  and  $C = -A_-$ .

If  $A$  is a linear combination  $b\rho + c\sigma$  of two density operators, define  $\mu(A) = b\mu(\rho) + c\mu(\sigma)$ . Using the convex linearity of  $\mu$  on the density operators, it is easy to check that if  $A$  has another such representation  $b'\rho' + c'\sigma'$  then  $b\mu(\rho) + c\mu(\sigma) = b'\mu(\rho') + c'\mu(\sigma')$  which means that  $\mu(A)$  is well-defined.

The uniqueness of the extension is obvious. It remains to check that the extended  $\mu$  is bounded. In fact, we show more, namely that i.e.  $\|\mu(A)\| \leq 1$  if  $\|A\| \leq 1$ . So let  $A \in \mathcal{T}$  and  $\|A\| \leq 1$ . As we saw above, there are positive trace-class operators  $B, C$  such that  $A = B - C$ . Then  $A = \|B\|\rho - \|C\|\sigma$  for some density operators  $\rho, \sigma$  where  $b, c \geq 0$  and  $b + c = \|A\| \leq 1$ . Now,  $\mu(\rho)$  and  $\mu(\sigma)$  are measures with norm 1. So  $\|\mu(A)\| \leq b\mu(\rho) + c\mu(\sigma) \leq b + c \leq 1$ .  $\square$

Let  $\mathcal{M}'$  be the space of the functionals  $\mathcal{M} \rightarrow \mathbb{R}$  where  $\mathbb{R}$  is the set of real numbers. Similarly let  $\mathcal{T}'$  be the space of the functionals  $\mathcal{T} \rightarrow \mathbb{R}$ .  $\mu$  gives rise to a dual transformation  $\mu' : \mathcal{M}' \rightarrow \mathcal{T}'$  that sends any  $h \in \mathcal{M}'$  to  $\mu'(h) = h \circ \mu$  so that

$$\mu'(h)(A) = h(\mu(A)) \quad \text{for all } h \in \mathcal{M}' \text{ and all } A \in \mathcal{T}. \quad (2)$$

Every measurable function  $f \in \mathcal{F}$  induces a functional  $\bar{f} \in \mathcal{M}'$  by integration:  $\bar{f}(\mu) = \int_{\Lambda} f d\mu$ . This gives rise to a transformation  $\nu : \mathcal{F} \rightarrow \mathcal{T}'$  that sends every  $f$  to  $\mu'(\bar{f})$ . Specifying  $h$  to  $\bar{f}$  in Equation 2 gives

$$(\nu f)(A) = \int_{\Lambda} f d\mu(A) \quad \text{for all } f \in \mathcal{F} \text{ and all } A \in \mathcal{T}. \quad (3)$$

Here and below we omit the parentheses around the argument of  $\nu$ .

**Lemma 5.** *For every  $f \in \mathcal{F}$ , there is a unique  $B \in \mathcal{B}$  with  $(\nu f)(A) = \text{Tr}(B \cdot \rho)$  for all density operators  $\rho$ .*

*Proof of Lemma 5.* Every  $B \in \mathcal{B}$  induces a functional  $\bar{B} \in \mathcal{T}'$  by  $\bar{B}(A) = \text{Tr}(B \cdot A)$ . Here  $\text{Tr}(B \cdot \rho)$  is well-defined because the product of a bounded operator and a trace-class operator is again in the trace class [11, Lemma 3, p. 38].

The map  $B \mapsto \bar{B}$  is an isometric isomorphism between  $\mathcal{B}$  and  $\mathcal{T}'$  [11, Theorem 2, p. 47]. So, for every  $X \in \mathcal{T}'$ , there is a unique  $B_X \in \mathcal{B}$  such that  $X(A) = \text{Tr}(B_X \cdot A)$  for all  $A \in \mathcal{T}$ . Furthermore, there is a unique  $B_X \in \mathcal{B}$  such that  $X(\rho) = \text{Tr}(B_X \cdot \rho)$  for all density operators  $\rho$ . This is because, as we showed above, the linear span of the density matrices is the whole space  $\mathcal{T}$ . The lemma follows because every  $\nu f$  belongs to  $\mathcal{T}'$ .  $\square$

For any  $f \in \mathcal{F}$ , the unique operator  $B$  with  $(\mathbf{v}f)(A) = \text{Tr}(B \cdot \rho)$  for all  $\rho$  will be denoted  $[\mathbf{v}f]$ .

**Lemma 6.**  $[\mathbf{v}F(E)] = E$  for every rank-1 projection  $E$ , and  $[\mathbf{v}1] = I$  where  $I$  is the constant function with value 1 and  $I$  is the unit matrix.

*Proof of Lemma 6.* Lemma 5 and equation (3) give

$$\text{Tr}([\mathbf{v}f] \cdot \rho) = \int_{\Lambda} f d\mu(\rho) \quad \text{for every density operator } \rho. \quad (4)$$

Equations (1) and (4) imply

$$\text{Tr}(\rho \cdot E) = \int_{\Lambda} f_E d\mu(\rho) = \text{Tr}([\mathbf{v}F(E)] \cdot \rho) \quad \text{for every density operator } \rho. \quad (5)$$

The right sides of Equations (1) and (4) coincide if we specify  $A$  to  $\rho$  and  $g$  to  $f_E$ . Therefore their left sides are equal.

$$\text{Tr}(E\rho) = \text{Tr}([\mathbf{v}f_E]\rho) \quad (6)$$

We now invoke the last clause in Definition 1 to find that, for all rank-1 projections  $E$  and all density matrices  $\rho$ ,

$$\text{Tr}(E\rho) = \int_{\Lambda} F(E) d\mu(\rho) = \text{Tr}([\mathbf{v}F(E)]\rho).$$

But this is, as we saw in the proof of Lemma 5, enough to show that  $[\mathbf{v}F(E)] = E$ .

By Lemma 5, we see that  $\mu'(1)$  is the unique operator that satisfies, for all  $\rho$ ,

$$\text{Tr}([\mathbf{v}1]\rho) = \int_{\Lambda} d\mu(\rho) = (\mu(\rho))(\Lambda) = 1 = \text{Tr}(\rho) = \text{Tr}(I\rho),$$

where the third equality comes from the fact that  $\mu$  maps density matrices to probability measures. Thus,  $[\mathbf{v}1] = I$ .  $\square$

**Lemma 7.** For any two rank-1 projections  $A, B$  of  $\mathcal{H}$ , there exists an operator  $H \in \mathcal{B}$  such that all four of  $H, A - H, B - H$ , and  $I - A - B + H$  are positive operators.

*Proof of Lemma 7.* Recall that an operator  $A$  is said to be positive if  $\langle \psi | A | \psi \rangle \geq 0$  for all  $|\psi\rangle \in \mathcal{H}$  and that  $A \leq B$  means that  $B - A$  is positive. A function  $f \in \mathcal{F}$  is *nonnegative* if  $f(\lambda) \geq 0$  for all  $\lambda \in \Lambda$ .

**Claim 8.** If  $f \in \mathcal{F}$  is nonnegative then  $[\mathbf{v}f]$  is a positive operator. Therefore, if  $f \leq g$  pointwise in  $\mathcal{F}$  then  $[\mathbf{v}f] \leq [\mathbf{v}g]$  in  $\mathcal{B}$ .

*Proof of Claim 8.* The second assertion follows immediately from the first applied to  $g - f$ , because  $\mathbf{v}$  is linear. To prove the first assertion, suppose  $f \in \mathcal{F}$  is nonnegative, and let  $|\psi\rangle$  be any vector in  $\mathcal{H}$ . The conclusion we want to deduce,  $\langle \psi | [\mathbf{v}f] | \psi \rangle \geq 0$ , is obvious if  $|\psi\rangle = 0$ , so we may assume that  $|\psi\rangle$  is a non-zero vector. Normalizing it, we may assume further that its length is 1. Then  $|\psi\rangle\langle\psi|$  is a density operator and therefore  $\mu(|\psi\rangle\langle\psi|)$  is a measure. Using equation (6), we compute

$$\langle \psi | [\mathbf{v}f] | \psi \rangle = \text{Tr}([\mathbf{v}f] |\psi\rangle\langle\psi|) = \int_{\Lambda} f d\mu(|\psi\rangle\langle\psi|) \geq 0,$$

where we have used that both the measure  $\mu(|\psi\rangle\langle\psi|)$  and the integrand  $f$  are nonnegative.  $\square$

Let  $\mathcal{F}_{[0,1]}$  be the subset of  $\mathcal{F}$  comprising the functions all of whose values are in the interval  $[0, 1]$ .

**Claim 9.** For any  $f, g \in \mathcal{F}_{[0,1]}$  there exists  $h \in \mathcal{F}_{[0,1]}$  such that all four of  $h$ ,  $f - h$ ,  $g - h$ , and  $1 - f - g + h$  are nonnegative.

*Proof of Claim 9.* Define  $h(\lambda) = \min\{f(\lambda), g(\lambda)\}$  for all  $\lambda \in \Lambda$ . Then the first three of the assertions in the lemma are obvious, and the fourth becomes obvious if we observe that  $f + g - h = \max\{f, g\} \leq 1$ .  $\square$

Now we are ready to complete the proof of Lemma 7. Apply Lemma 9 with  $f = \mu(A)$  and  $g = \mu(B)$ , let  $h$  be the function given by the lemma, and let  $H = v(h)$ . The nonnegativity of  $h$ ,  $f - h$ ,  $g - h$ , and  $1 - f - g + h$  implies, by Claim 8, the positivity of  $v(h) = H$ ,  $v(S(A) - h) = A - H$ ,  $v(S(B) - h) = B - H$ , and  $v(1 - F(A) - F(B) + h) = I - A - B + H$ , where we have also used the linearity of  $v$ , the fact that  $v(1) = I$ , and the formula  $v(\mu(A)) = A$  for all  $A$  in the domain of  $\mu$ .  $\square$

Now we are ready to prove Theorem 3. Let us apply Lemma 6 to two specific rank-1 projections. Fix two orthonormal vectors  $|0\rangle$  and  $|1\rangle$ . (This is where we use that  $\mathcal{H}$  has dimension at least 2.) Let  $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ . We use the projections  $A = |0\rangle\langle 0|$  and  $B = |+\rangle\langle +|$  to the subspaces spanned by  $|0\rangle$  and  $|+\rangle$ . Let  $H$  be as in Lemma 7 for these projections  $A$  and  $B$ .

From the positivity of  $H$  and of  $A - H$ , we get that  $0 \leq \langle 1|H|1\rangle$  and that

$$0 \leq \langle 1|(A - H)|1\rangle = \langle 1|A|1\rangle - \langle 1|H|1\rangle = -\langle 1|H|1\rangle,$$

where we have used that  $|1\rangle$ , being orthogonal to  $|0\rangle$ , is annihilated by  $A$ . Combining the two inequalities, we infer that  $\langle 1|H|1\rangle = 0$  and therefore, since  $H$  is positive,  $H|1\rangle = 0$ . Similarly, using the orthogonal vectors  $|+\rangle$  and  $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$  in place of  $|0\rangle$  and  $|1\rangle$ , we obtain  $H|-\rangle = 0$ . So, being linear,  $H$  is identically zero on the subspace of  $\mathcal{H}$  spanned by  $|1\rangle$  and  $|-\rangle$ ; note that  $|0\rangle$  is in this subspace, so we have  $H|0\rangle = 0$ .

Now we use the positivity of  $I - A - B + H$ . Since  $H|0\rangle = 0$ , we can compute

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

This contradiction completes the proof of the theorem.  $\square$

**Remark 10** (Symmetry or the lack of thereof). In view of the idea of symmetry or even-handedness suggested by Spekkens [12], one might ask whether there is a dual version of Theorem 3, that is, a version that requires convex-linearity for effects but looks only at pure states and does not require any convex-linearity for states. The answer is no; with such requirements there is a trivial example of a successful hidden-variable theory, regardless of the dimension of the Hilbert space. The theory can be concisely described as taking the quantum state itself as the ‘‘hidden’’ variable. In more detail, let  $\Lambda$  be the set of all pure states. Let  $\mu$  assign to each operator  $|\psi\rangle\langle\psi|$  the probability measure on  $\Lambda$  concentrated at the point  $\lambda_{|\psi\rangle}$  that corresponds to the vector  $|\psi\rangle$ . Let  $F$  assign to each effect  $E$  the function on  $\Lambda$  defined by

$$F(E)(\lambda_{|\psi\rangle}) = \langle\psi|E|\psi\rangle.$$

We have trivially arranged for this to give the correct expectation for any effect  $E$  and any pure state  $|\psi\rangle$ . The formula for  $F(E)$  is clearly convex-linear (in fact, linear) as a function of  $E$ . Of course,  $\mu$  cannot be extended convex-linearly to mixed states, so that Theorem 3 does not apply.

### 3 Value No-Go Theorems

Value no-go theorems assert that hidden-variable theories cannot even produce the correct outcomes for individual measurements, let alone the correct probabilities or expectation values. Such theorems considerably predated the expectation no-go theorems considered in the preceding section. Value no-go theorems were first established by Bell [1, 2] and then by Kochen and Specker [9]; we shall also refer to the user-friendly exposition given by Mermin [10]. To formulate value no-go theorems, one must specify what “correct outcomes for individual measurements” means.

**Definition 11.** *Let  $\mathcal{H}$  be a Hilbert space, and let  $\mathcal{O}$  be a set of observables, i.e., self-adjoint operators on  $\mathcal{H}$ . A valuation for  $\mathcal{O}$  in  $\mathcal{H}$  is a function  $v$  assigning to each observable  $A \in \mathcal{O}$  a number  $v(A)$  in the spectrum 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.*

The intention behind this definition is that, in a hidden-variable theory, a quantum state represents an ensemble of individual systems, each of which has definite values for observables. That is, each individual system has a valuation associated to it, describing what values would be obtained if we were to measure observable properties of the system. A believer in such a hidden-variable theory would expect a valuation for the set of all self-adjoint operators on  $\mathcal{H}$ , unless there were superselection rules rendering some such operators unobservable.

Before we proceed, we recall the notion of joint spectra [3, Section 6.5].

**Definition 12.** *The joint spectrum  $\sigma(A_1, \dots, A_n)$  of pairwise commuting, self-adjoint operators  $A_1, \dots, A_n$  on a Hilbert space  $\mathcal{H}$  is a subset of  $\mathbb{R}^n$ . If  $A_1, \dots, A_n$  are simultaneously diagonalizable then  $(\lambda_1, \dots, \lambda_n) \in \sigma(A_1, \dots, A_n)$  iff there is a non-zero vector  $|\psi\rangle$  with  $A_i|\psi\rangle = \lambda_i|\psi\rangle$  for  $i = 1, \dots, n$ . In general,  $(\lambda_1, \dots, \lambda_n) \in \sigma(A_1, \dots, A_n)$  iff for every  $\varepsilon > 0$  there is a unit vector  $|\psi\rangle \in \mathcal{H}$  with  $\|A_i|\psi\rangle - \lambda_i|\psi\rangle\| < \varepsilon$  for  $i = 1, \dots, n$ .*

**Proposition 13.** *For any continuous function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f(A_1, \dots, A_n) = 0$  if and only if  $f$  vanishes identically on  $\sigma(A_1, \dots, A_n)$ .*

The proposition is implicit in the statement, on page 155 of [3], that “most of Section 1, Subsection 4, about functions of a one operator,” can be repeated in the context of several commuting operators. We give a detailed proof of the proposition in [4, §4.1].

**Theorem 14** ([2, 9, 10]). *If  $\text{Dim}(\mathcal{H}) = 3$  then there is a finite set  $\mathcal{O}$  of rank 1 projections for which no valuation exists.*

The proof of Theorem 14 can be derived from the work of Bell [2, Section 5], and we do that explicitly in [4, §4.3]. The construction given by Kochen and Specker [9] provides the desired  $\mathcal{O}$  more directly. The proof of Theorem 1 in [9] uses a Boolean algebra generated by a finite set of one-dimensional subspaces of  $\mathcal{H}$ , and it shows that the projections to those subspaces constitute an  $\mathcal{O}$  of the required sort. Mermin’s elegant exposition [10, Section IV] deals instead with squares  $S_i^2$  of certain spin-components of a spin-1 particle, but these are projections to 2-dimensional subspaces of  $\mathcal{H}$ , and the complementary rank-1 projections  $I - S_i^2$  serve as the desired  $\mathcal{O}$ .

**Theorem 15** (Second Bootstrapping Theorem). *Suppose  $\mathcal{H} \subseteq \mathcal{H}'$  are finite-dimensional Hilbert spaces. Suppose further that  $\mathcal{O}$  is a finite set of rank-1 projections of  $\mathcal{H}$  for which no valuation exists. Then there is a finite set  $\mathcal{O}'$  of rank-1 projections of  $\mathcal{H}'$  for which no valuation exists.*

This is our second bootstrapping theorem. Intuitively, such dimension bootstrapping results are to be expected. If hidden-variable theories could explain the behavior of quantum systems described by

the larger Hilbert space, say  $\mathcal{H}'$ , then they could also provide an explanation for systems described by the subspace  $\mathcal{H}$ . The latter systems are, after all, just a special case of the former, consisting of the pure states that happen to lie in  $\mathcal{H}$  or mixtures of such states. But often no-go theorems give much more information than just the impossibility of matching the predictions of quantum-mechanics with a hidden-variable theory. They establish that hidden-variable theories must fail in very specific ways. It is not so obvious that these specific sorts of failures, once established for a Hilbert space  $\mathcal{H}$ , necessarily also apply to its superspaces  $\mathcal{H}'$ .

*Proof.* Clearly, if two Hilbert spaces are isomorphic and if one of them has a finite set  $\mathcal{O}$  of rank-1 projections with no valuation, then the other also has such a set. It suffices to conjugate the projections in  $\mathcal{O}$  by any isomorphism between the two spaces. Thus, the existence of such a set  $\mathcal{O}$  depends only on the dimension of the Hilbert space, not on the specific space.

Proceeding by induction on the dimension of  $\mathcal{H}'$ , we see that it suffices to prove the theorem in the case where  $\dim(\mathcal{H}') = \dim(\mathcal{H}) + 1$ . Given such  $\mathcal{H}$  and  $\mathcal{H}'$ , 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  $\mathcal{H}$  and thus isomorphic to  $\mathcal{H}$ . By the induction hypothesis, this subspace  $|\psi\rangle^\perp$  has a finite set  $\mathcal{O}$  of rank-1 projections for which no valuation exists. Each element of  $\mathcal{O}$  can be regarded as a rank-1 projection of  $\mathcal{H}'$ ; indeed, if the projection was given by  $|\phi\rangle\langle\phi|$  in  $|\psi\rangle^\perp$ , then we can just interpret the same formula  $|\phi\rangle\langle\phi|$  in  $\mathcal{H}'$ , using the same unit vector  $|\phi\rangle \in |\psi\rangle^\perp$ .

Let  $\mathcal{O}_1$  consist of all the projections from  $\mathcal{O}$ , interpreted as projections of  $\mathcal{H}'$ , together with one additional rank-1 projection, namely  $|\psi\rangle\langle\psi|$ . What can a valuation  $v$  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 valuation for  $\mathcal{O}$ . But  $\mathcal{O}$  was chosen so that it has no valuation, and so we cannot have  $v(|\psi\rangle\langle\psi|) = 0$ . Therefore  $v(|\psi\rangle\langle\psi|) = 1$ . (It follows that  $v$  maps the projections associated to all the other elements of  $\mathcal{O}'$  to zero, but we shall not need this fact.)

We have thus shown that any valuation 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 projections such that any valuation for  $\mathcal{O}_2$  must send  $|\psi'\rangle\langle\psi'|$  to 1. No valuation 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 valuation for the union  $\mathcal{O}_1 \cup \mathcal{O}_2$ , which thus serves as the  $\mathcal{O}'$  required by the theorem.  $\square$

**Theorem 16** (Value no-go theorem). *Suppose that the dimension of the Hilbert space is at least 3.*

1. *There is a finite set  $\mathcal{O}$  of projections for which no valuation exists.*
2. *If the dimension is finite then there is a finite set  $\mathcal{O}$  of rank 1 projections for which no valuation exists.*

The desired finite sets of projections are constructed explicitly in the proof. The finiteness assumption in part (2) of the theorem cannot be omitted. If  $\dim(\mathcal{H})$  is infinite, then the set  $\mathcal{O}$  of all finite-rank projections admits a valuation, namely the constant zero function. This works because the definition of “valuation” imposes constraints on only finitely many observables at a time.

*Proof.* When the dimension of  $\mathcal{H}$  is greater than 3, but still finite, we use our Second Bootstrapping Theorem. Notice that, if one merely wants a no-go theorem saying that some  $\mathcal{O}$  has no valuation, then this bootstrapping is easy, as noted in [1, 9, 10]. Work is needed only to get all the operators in  $\mathcal{O}$  to be rank 1 projections.



It remains to treat the case of infinite-dimensional  $\mathcal{H}$ . Let  $\mathcal{K}$  and  $\mathcal{L}$  be Hilbert spaces, with  $\dim(\mathcal{K}) = 3$  and  $\dim(\mathcal{L}) = \dim(\mathcal{H})$ . Note that then their tensor product  $\mathcal{K} \otimes \mathcal{L}$  has the same dimension as  $\mathcal{H}$ , so it can be identified with  $\mathcal{H}$ .

Let  $\mathcal{O}$  be as in Theorem 14 for the 3-dimensional  $\mathcal{K}$ . Let  $\mathcal{O}' = \{P \otimes I_{\mathcal{L}} : P \in \mathcal{O}\}$ , where  $I_{\mathcal{L}}$  is the identity operator on  $\mathcal{L}$ . Then  $\mathcal{O}'$  is a set of infinite-rank projections of  $\mathcal{K} \otimes \mathcal{L} = \mathcal{H}$ , having the same algebraic structure as  $\mathcal{O}$ . It follows that there is no valuation for  $\mathcal{O}'$ .  $\square$

Let's say that a projection  $A$  on Hilbert space  $\mathcal{H}$  is a *rank- $n$  projection modulo identity* if either  $A$  is of rank  $n$  or else  $\mathcal{H}$  splits into a tensor product  $\mathcal{K} \otimes \mathcal{L}$  such that  $\mathcal{K}$  is finite-dimensional and  $A$  has the form  $P \otimes I_{\mathcal{L}}$  where  $P$  is of rank  $n$  and  $I_{\mathcal{L}}$  is the identity operator on  $\mathcal{L}$ . The proof of Theorem 16 gives us the following corollary.

**Corollary 17.** *If the dimension of the Hilbert space is at least 3 then there is a finite set of rank-1 projections modulo identity for which no valuation exists.*

## 4 One successful hidden-variable theory

By reducing both species of no-go theorems to projection measurement, where measurement as observable and measurement as effect coincide, we made it easier to see similarities and differences. No, the expectation no-go theorem does not imply the value no-go theorem. But the task of proving this claim formally, say for a given dimension  $d = \text{Dim}(\mathcal{H})$ , is rather thankless. You have to construct a counterfactual physical world where the expectation no-go theorem holds but the value no-go theorem fails. There is, however, one exceptional case, that of dimension 2. Theorem 3 assumes  $\text{Dim}(\mathcal{H}) \geq 2$  while Theorem 16 assumes  $\text{Dim}(\mathcal{H}) \geq 3$ . So what about dimension 2?

Bell developed, in [1] and [2], a hidden-variable theory for a two-dimensional Hilbert space  $\mathcal{H}$ . Here we summarize the improved version of Bell's theory due to Mermin [10], we simplify part of Mermin's argument, and we explain why the theory doesn't contradict Theorem 3.

In the rest of this section, we work in the two-dimensional Hilbert space  $\mathcal{H}$ . Let  $\mathcal{V}$  be the set of value maps  $v$  for all the observables on  $\mathcal{H}$ . In each pure state  $\psi$ , the hidden variables should determine a particular member of  $V$ .

**Definition 18.** A *value representation* for quantum systems described by  $\mathcal{H}$  is a pair  $(\Lambda, V)$  where

- $\Lambda$  is a probability space and
- $V$  a function  $\psi \rightarrow V_{\psi}$  on the pure states such that every  $V_{\psi}$  is a map  $\lambda \rightarrow V_{\psi}^{\lambda}$  from (the sample space of)  $\Lambda$  onto  $\mathcal{V}$ .

Further, we require that, for any pure state  $\psi$  and any observable  $A$ , the expectation  $\int_{\Lambda} V_{\psi}^{\lambda}(A) d\lambda$  of the eigenvalue of  $A$  agrees with the prediction  $\langle \psi | A | \psi \rangle$  of quantum theory:

$$\int_{\Lambda} V_{\psi}^{\lambda}(A) d\lambda = \langle \psi | A | \psi \rangle \quad (7)$$

Definition 18 is narrowly tailored for our goals in this section; in the full paper we will give a general definition of value representation. Notice that, if a random variable (in our case, the eigenvalue of  $A$  in  $\psi$ ) takes only two values, then the expected value determines the probability distribution. A priori we should be speaking about commuting operators and joint spectra but things trivialize in the 2-dimensional case. Recall Proposition 13 and notice that, in the 2-dimensional Hilbert space, if operators  $A, B$  commute, then one of them is a polynomial function of the other.

**Theorem 19.** *There exists a value representation for the quantum systems described by the two-dimensional Hilbert system  $\mathcal{H}$ .*

*Proof.* Let  $\vec{\sigma}$  be the triple of the Pauli matrices  $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ,  $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ ,  $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ . For any  $\vec{n} \in \mathbb{R}^3$ , the dot product  $\vec{n} \cdot \vec{\sigma}$  is a Hermitian operator with eigenvalues  $\pm 1$ . Every pure state of  $\mathcal{H}$  is an eigenstate, for eigenvalue  $+1$ , of  $\vec{n} \cdot \vec{\sigma}$  for a unique  $\vec{n}$ . We use the notation  $|\vec{n}\rangle$  for this eigenstate.

If  $\mathcal{H}$  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  $\vec{n}$ . It is a special property of spin  $\frac{1}{2}$  that all pure states are of this form; for higher spins, a superposition of states with definite spin directions need not have a definite spin direction.

On  $\mathcal{H}$ , any Hermitian operator  $A$  has the form  $a_0I + \vec{a} \cdot \vec{\sigma}$  for some scalar  $a_0 \in \mathbb{R}$  and vector  $\vec{a} \in \mathbb{R}^3$ . The eigenvalues of  $A$  are  $a_0 \pm \|\vec{a}\|$ . Observables  $a_0I + \vec{a} \cdot \vec{\sigma}$  and  $b_0I + \vec{b} \cdot \vec{\sigma}$  commute if and only if  $\vec{a}$  and  $\vec{b}$  are either parallel or antiparallel.

The desired probability space is the set  $S^2$  of unit vectors in  $\mathbb{R}^3$  with the uniform probability measure. Let  $\vec{m}$  range over  $S^2$ . Then

$$V_{\vec{n}}^{\vec{m}}(a_0I + \vec{a} \cdot \vec{\sigma}) = \begin{cases} a_0 + \|\vec{a}\| & \text{if } (\vec{m} + \vec{n}) \cdot \vec{a} \geq 0, \\ a_0 - \|\vec{a}\| & \text{if } (\vec{m} + \vec{n}) \cdot \vec{a} < 0. \end{cases}$$

It remains to check that

$$\int_{S^2} V_{\vec{n}}^{\vec{m}}(a_0I + \vec{a} \cdot \vec{\sigma}) d\vec{m} = \langle \vec{n} | (a_0I + \vec{a} \cdot \vec{\sigma}) | \vec{n} \rangle. \quad (8)$$

We begin with a couple of simplifications. First, we may assume that  $a_0 = 0$ , because a general  $a_0$  would just be added to both sides of Equation (8). Second, thanks to the rotational symmetry of the situation (where rotations are applied to all three of  $\vec{a}$ ,  $\vec{n}$  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 that the right side of Equation (8) is  $n_z$ .

So our task is to prove that the average over  $\vec{m}$  of the values assigned to  $\sigma_z$  is  $n_z$ . By definition, the value assigned to  $\sigma_z$  is  $\pm 1$ , where the sign is chosen to agree with that of  $m_z + n_z$ . In view of how  $\vec{m}$  is chosen, 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  $S^2$  horizontally at a level  $n_z$  below the sphere's center. By a theorem of Archimedes, when a sphere is cut by a plane, its area is divided in the same ratio as the length of the diameter perpendicular to the plane. So the 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.  $\square$

Finally, we explain why Bell's theory doesn't contradict Theorem 3. To obtain an expectation representation, we must extend the map  $V$  convex-linearly to all density matrices. But 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 projections

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

respectively. Averaging the first two with equal weights, we get  $\frac{1}{2}I$ ; averaging the last two gives the same result. So a convex-linear extension  $T$  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 concentrated on the union of two unit spheres tangent to the  $y$ - $z$ -plane at the origin, while the second is concentrated on the union of two unit spheres tangent to the  $x$ - $y$ -plane at the origin.

Thus, Bell's hidden-variable theory for 2-dimensional  $\mathcal{H}$  does not clash with the fact that Theorem 3 applies in the 2-dimensional case. To obtain an expectation representation, we must extend this map convex-linearly to all density matrices.

Thus, Bell's example of a hidden-variable theory for 2-dimensional  $\mathcal{H}$  does not fit the assumptions in any of the expectation no-go theorems. It does not, therefore, clash with the fact that those theorems, unlike the value no-go theorems, apply in the 2-dimensional case.

## References

- [1] John S. Bell (1964): *On the Einstein-Podolsky-Rosen paradox*. *Physics* 1, pp. 195–200.
- [2] John S. Bell (1966): *On the problem of hidden variables in quantum mechanics*. *Reviews of Modern Physics* 38, pp. 447–452.
- [3] Michael S. Birman & Michael Z. Solomjak (1987): *Spectral Theory of Self-Adjoint Operators in Hilbert Space*. Reidel. Originally in Russian, Leningrad University Press, 1980.
- [4] Andreas Blass & Yuri Gurevich (Sept. 2015): *On Hidden Variables: Value and Expectation No-Go Theorems*. arXiv:1509.06896.
- [5] Sławomir Bugajski (1993): *Classical frames for a quantum theory—a bird's-eye view*. *International Journal of Theoretical Physics* 32, pp. 969–977.
- [6] Christopher Ferrie & Joseph Emerson (2008): *Frame representations of quantum mechanics and the necessity of negativity in quasi-probability representations*. *Journal of Physics A: Mathematical and Theoretical* 41:352001. See also arXiv:0711.2658.
- [7] Christopher Ferrie & Joseph Emerson (2009): *Framed Hilbert space: hanging the quasi-probability pictures of quantum theory*. *New Journal of Physics* 11:063040. See also arXiv:0903.4843.
- [8] Christopher Ferrie, Ryan Morris & Joseph Emerson (2010): *Necessity of negativity in quantum theory*. *Physical Review A* 82:044103. See also arXiv:0910.3198.
- [9] Simon Kochen & Ernst Specker (1967): *The problem of hidden variables in quantum mechanics*. *Journal of Mathematics and Mechanics* 17, pp. 59–87.
- [10] N. David Mermin (1993): *Hidden variables and the two theorems of John Bell*. *Reviews of Modern Physics* 65, pp. 803–815.
- [11] Robert Schatten (1970): *Norm ideals of completely continuous operators*. Springer Verlag. 2nd edition.
- [12] Robert W. Spekkens (2008): *Negativity and contextuality are equivalent notions of nonclassicality*. *Physics Review Letters* 101(2):020401. See also arXiv:0710.5549.