Contextuality lecture notes

Matt Pusey

June 20, 2019

Part I
The Kochen-Specker theorem

1 Mathematical statement

Given a set S of rank 1 projectors, a function $v : S \mapsto \{0, 1\}$ is called a *valid value assignment* if $\sum_{P \in M} v(P) = 1$ whenever $M \subseteq S$ is a resolution of the identity, i.e. $\sum_{P \in M} P = I$.

If no valid value assignment exists, we call S a *Kochen-Specker (KS) set.*

Theorem 1 (Kochen-Specker [1])**.** *Finite KS sets exist in all Hilbert spaces of dimension greater than* 2*.*

2 Physics interlude

Sets of projectors M with $\sum_{P \in M} P = I$ represent projective measurements. We can think of a valid value assignment as pre-determining outcomes for all of the measurements $M \subseteq S$, by taking the $P \in M$ with $v(P) = 1$ to be the outcome that will occur. The validity condition that $\sum_{P \in M} v(P) = 1$ simply ensures there is exactly one such P, so that we don't have multiple outcomes occurring, or none at all. But non-trivial physical assumptions are baked into the very notion of a value assignment:

1. *Outcome determinism:* There is a fact about which outcome will occur, so that the apparent randomness of quantum measurements is purely due to ignorance of the true value assignment.

2. *Measurement non-contextuality:* Whether or not an outcome occurs is determined solely by the projector representing it, and no other information (or "context"). In particular, whether it occurs is independent

of the measurement it appears in, so that we write $v(P)$ instead of $v(P, M)$.

To put it another way, the assumption is that the projectors in S represent objective properties of a system, and a measurement simply reveals which of some set of mutually exclusive and exhaustive properties holds. [These assumptions will be discussed more in part II.]

Both assumptions are necessary. Without determinism, non-contextual probability assignments exist, for example the operational probabilities (given in quantum theory by the Born rule). Without non-contextuality, one can assign an outcome to every measurement arbitrarily, since there are no longer any consistency requirements between measurements.

3 Some useful reformulations

3.1 Higher dimensions are free

Suppose we have found a Kochen-Specker set S in dimension d. We can easily construction a Kochen-Specker set S' in dimension $d+1$ as follows. Pick a basis $\{|0\rangle, \ldots, |d\rangle\}$ for the $d+1$-dimensional space. The subspace orthogonal to $|0\rangle$ has dimension d, so we can embed S into it as S_0. Similarly for the subspace orthogonal to $|1\rangle$ choose an embedding S_1. We define $S' := S_0 \cup S_1 \cup M_b$ as those two embeddings along with the basis projectors $M_b := \{|0\rangle\langle 0|, \ldots, |d\rangle\langle d|\}$.

Suppose we have a valid value assignment $v' : S' \mapsto \{0, 1\}$. Since M_b sums to identity, we either have $v'(|0\rangle\langle 0|) = 0$ or $v'(|1\rangle\langle 1| = 0$. Suppose the first condition holds. If $M \subset S$ sums to identity in dimension d, then corresponding subset $M' \subset S_0$ of the first embedding, along with $|0\rangle\langle 0|$, sum to identity in dimension $d+1$. Since v' is valid, $\sum_{P \in M'} v'(P) + v'(|0\rangle\langle 0|) = 1$. But $v'(|0\rangle\langle 0|) = 0$ so we have $\sum_{P \in M'} v'(P) = 1$. Hence we obtain a valid value assignment $v : S \mapsto \{0, 1\}$ by letting $v(P)$ equal $v'(P')$ where $P' \in S_0$ is the embedding of P.

If $v'(|1\rangle\langle 1|) = 0$ we can run the same argument on the second embedding. Either way, we end up with a valid value assignment on S. So if no such assignment exists, we have a contradiction and so the value assignment to S' we have been supposing cannot exist.

The upshot is that the Kochen-Specker theorem can be proven simply by exhibiting a KS set in dimension 3.

3.2 Can use observables

Suppose we have a set of observables (Hermitian operators) O. Some subsets of the observables will commute and hence have basis in which they are jointly diagonalisable. Let S consist projectors onto the elements of all such bases. Then a valid value assignment $v : S \to \{0,1\}$ induces an assignment $v : O \to \mathbb{R}$ where $f(A, B, \dots) = 0$ for some commuting $A, B, \dots \in O$ then $f(v(A), v(B), \dots) = 0$. Important implications of this property include:

1. By taking f to be the characteristic polynomial of A, we see that $v(A)$ is an eigenvalue of A for all $A \in O$.

2. By taking f to be $f(A, B, C) = A+B-C$ we see that if $A, B, A+B \in O$ and A commutes with B then $v(A+B) = v(A) + v(B)$.

3. By taking f to be $f(A, B, C) = AB - C$ we see that if $A, B, AB \in O$ (which by the Hermiticity of AB implies that A and B commute) we have $v(AB) = v(A)v(B)$.

The assignment $v : (A)$ for $A \in O$ is defined as follows. Write the spectral decomposition of A as $A = \sum_i a_i \Pi_i$ where Π_i are orthogonal (not necessarily rank 1) projectors summing to I, and the a_i are distinct eigenvalues of A. By construction of S it contains projectors onto at least one eigenbasis of A and so we can write $\Pi_i = \sum_j P_{ij}$ where P_{ij} are orthogonal elements of S. We set $v(A) = \sum_{i,j} a_i v(P_{ij})$. But what we can also write some $\Pi_i = \sum_j Q_{ij}$ for some other $Q_{ij} \in S$? Since S contains projectors onto an at least one eigenbasis of A, it contains projectors R_k such that $Pi_i + \sum_k R_k = I$. But then since $v : S \to \{0,1\}$ is valid, $\sum_j v(P_{ij}) + \sum_j v(R_k) = 1 = \sum_j v(Q_{ij}) + \sum_j v(R_k)$ and so $\sum_j v(P_{ij}) = \sum_j v(Q_{ij})$ and hence either contribute the same to $v(O)$.

Now suppose that $A, B, \dots \in O$ commute. By the definition of S, projectors P_i onto some common eigenbasis are in S. Write $A = \sum_i a_i P_i$ (where now a_i can be repeated eigenvalues), $B = \sum_i b_i P_i$ etc. $f(A, B, \dots) = 0$ simply means that $f(a_i, b_i, \dots) = 0$ for all i. By the above $v(A) = \sum_i a_i v(P_i) = a_j$ where j is the unique index for which $v(P_j) = 1$ (with all other $v(P_i) = 0$). Similarly $v(B) = b_j$ etc. Hence we immediately obtain $f(v(A), v(B), \dots) = 0$.

The upshot is that the Kochen-Specker theorem can be proven by exhibiting some observables that cannot be given valid valuations in the above sense. This implicitly defines a KS set via the above construction.

4 The Mermin-Peres square

We will now use the reformulations of the previous section to prove that a finite KS set exists in all Hilbert spaces of dimension greater than 3. This is slightly weaker than Theorem 1 but is also somewhat easier to prove. From §3.1 we only need to exhibit a KS set in dimension 4. From §3.2 we can do this using observables. The observables O we use [2] are tensor products of Pauli observables, so that XY is shorthand for $\sigma_x \otimes \sigma_Y$ etc. It is useful to arrange them in a square:

$$\begin{array}{ccc} XI & IX & XX \\ IY & YI & YY \\ XY & YX & ZZ \end{array}.$$

Every row and column in this square has the property that the product of the first two elements is equal to the third, except the final column where we have $(XX)(YY) = -ZZ$. Suppose we had a valuation

$$\begin{array}{ccc} v(XI) & v(IX) & v(XX) \\ v(IY) & v(YI) & v(YY) \\ v(XY) & v(YX) & v(ZZ) \end{array}.$$

Since the eigenvalues of the observables are ± 1, the first numbered property in §3.2 tells us that the v must be ± 1. The third numbered property tells us that $v(XI)v(IY) = v(XY)$ etc, which we can also write as $v(XI)v(IY)v(XY) = 1$ etc. For the final column we have $v(XX)v(YY) = -v(ZZ)$[1] which we write $v(XX)v(YY)v(ZZ) = 1$. In other words we have a table of values ± 1 whose rows and columns all multiply to 1, except for the final column which multiples to -1. This means if we multiply the entire table row-wise we get 1, whereas column-wise we get -1. Since changing the order of multiplication for real numbers never changes the result, we have a contradiction.

5 Comparison with older no-go theorems

5.1 Von Neumann

Von Neumann [3] assumed that observables must be assigned values such that $v(A+B) = v(A)+v(B)$ regardless of whether A and B commute. This

[1]The eagle-eyed reader will notice this goes slightly beyond the third numbered property in §3.2, but it just as easily follows from the general condition involving $f(A, B, \dots)$.

is much stronger than assuming this condition only in the commuting case and gives a contradiction already in dimension 2.

5.2 Gleason

Gleason [4] showed that if, in dimension greater than 2, all the rank 1 projectors $\{P\}$ are assigned non-negative numbers $p(P)$ that sum to one over all resolutions of the identity, then $v(P) = \mathrm{Tr}(\rho P)$ for some positive operator ρ with $\mathrm{Tr}(\rho) = 1$. As noted by Bell [5], this in particular implies that $v(P) \notin \{0, 1\}$ for some P, so we get something like the Kochen-Specker theorem as a corollary. However the theorem needs an infinite set of projectors rather than the finite sets used in KS. Since we can only implement a finite number of measurements in an experiment, the KS theorem has more hope of an experimental counterpart.

6 Frameworks

KS sets and related structures can be studied using a variety of frameworks. This can be based on graph theory [6], hypergraph theory [7] (which makes the connection to Bell's theorem particular clear), sheaf theory [8] and more.

Part II
Operational contextuality

7 Two assumptions give rise to two questions

Bell's theorem can be considered a theoretical result that quantum theory is incompatible with local causality. However, we can go much further than this and do an experiment (a "Bell experiment") showing nature is incompatible with local causality, regardless of whether quantum theory is exactly true.

Can something similar be done, starting from the Kochen-Specker theorem? If so, this could give a route to experimentally certifying that single systems behave in a non-classical way (recall that Bell experiments require at least two systems).

The starting point for an experiment will presumably be to take some KS set an attempt to actually implement the measurements in it. Such an implementation will never be perfect, so the problem of operational contextuality is basically to have a principled way of saying how good the implementation has to be in order to say something interesting.

There are two important ways in which an implementation can deviate from the ideal, which are closely linked to the two "physical assumptions" discussed in §2:

1. The KS set is built from projective measurements. But an implementation will inevitably involve some POVMs that approximate those projective measurements.
2. The same projector appears in different measurements. But if we implement two different measurement procedures there will never be two outcomes that implement the exact projector (or POVM effect).

Therefore we need to know:

1. How close to projective measurement do we need to get?
2. How close to "the same projector" do we need to get?

Taking a step back, we really want to be able to analyse our experiment without reference to quantum theory. So we should be able say:

1. What does "projective measurement" even mean operationally?
2. What does "the same projector" even mean operationally?

These quantitive and qualitative questions are closely related. If the answers to the quantitive are stated in operational terms, then setting the "closeness" to zero will give answers to the qualitative questions. And if we can justify our answers to the qualitative questions by showing how they are connected to the impossibility of a non-classical model, we should be able to quantify this connection to answer the corresponding quantitive questions.

Two approaches to these questions have been taken, one is to start from the KS theorem and try to figure out good answers somehow (the "incremental approach"), the other is to give a totally operational notion of noncontextuality and work out the answers from that.

8 Comments on some incremental steps

The literature on testing the Kochen-Specker theorem without a wholesale new notion of noncontextuality is extensive but rather disjointed. I have not been able to identify a single fully worked-out proposal for how to do this. However there are several ideas that often appear. Here I evaluate whether those ideas make sense even within the setting of quantum theory.

8.1 Just use any measurement?

We might ask if the notion of noncontextuality used in the Kochen-Specker theorem can be directly generalised to arbitrary measurements. However, this doesn't give an interesting notion. Take the POVM (on a qubit, for example)

$$\left\{ \frac{I}{2}, \frac{I}{2} \right\}.$$

If we set $v(I/2) = 0$ then neither outcome occurs, and if we set $v(I/2) = 1$ both outcomes occur. Either is a violation of the condition that the values should sum to 1 for any measurement.

When we notice that this POVM can be implemented by throwing away the system and tossing an unrelated fair coin, we see that there is absolutely no reason to think of this measurement as revealing any pre-existing property of the system. Solving $v(I/2) + v(I/2) = 1$ gives $v(I/2) = 1/2$ which makes perfect sense, at least for this implementation. See [9] for further discussion.

8.2 Repeatability?

It has often been suggested that the crucial feature of projective measurements is that they are repeatable, i.e. if you do the same measurement again you get the same result.

However, this is not actually enough to ensure a measurement is projective. Consider the qutrit POVM

$$\left\{ |1\rangle\langle 1| + \frac{|2\rangle\langle 2|}{2}, \frac{|2\rangle\langle 2|}{2} + |1\rangle\langle 1| \right\}.$$

If the post-measurement state is $|1\rangle$ for the first outcome and $|3\rangle$ for the second, then this measurement is repeatable.

8.3 Sharpness

It has been shown [10] that projective measurements in quantum theory are exactly those with two operational properties:

1. Repeatability, as above.
2. Minimally disturbing: the measurement procedures only disturbs the statistics of incompatible measurements.

The fact this has been proven to be equivalent to projective measurements in quantum theory is a very good start. However, more needs to be done to show how the minimal disturbance condition can actually be checked using experimental data.

8.4 Boxes that measure observables

If we have a box designed to measure some observable, it seems reasonable to assume that it always does the same thing no matter where it in an experiment it is used. However, a proof of the Kochen-Specker theorem will now involve joint measurements of commuting observables. If we implement this by putting one box after another, we need to be sure that the value revealed by the second box is independent of whether the first box is there. Since we know quantum measurements are disturbing, this means we need to be sure the first measurement is compatible with the second. So the problem of identifying the same projector in different measurements has simply been shifted to the problem of identifying compatible measurement procedures.

9 Fully operational notion

A fully operational notion of noncontextuality has been proposed by Spekkens [11]. It uses the notion of an ontological model of an operational theory:

$$p(k|\mathcal{M},\mathcal{P}) = \int p(k|\mathcal{M},\lambda)p(\lambda|\mathcal{P})d\lambda$$

9.1 Measurement noncontextuality

We tackle the second question first, regarding "the same projectors" appearing in multiple measurements. In quantum theory two projectors are equal if and only if they get the same probability on all states:

$$P = P' \qquad \Longleftrightarrow \qquad \mathrm{Tr}(\rho P) = \mathrm{Tr}(\rho P')\ \forall \rho$$

The right hand side can be written operationally

$$p(k|\mathcal{M},\mathcal{P}) = p(k'|\mathcal{M}',\mathcal{P})\ \forall \mathcal{P}$$

In the Kochen-Specker theorem we assumed the same projector always gets the same value. But if we want to allow for noisy measurements we shouldn't assume outcome determinism. Hence we instead assume the probability for that outcome is the same:

$$p(k|\mathcal{M},\mathcal{P}) = p(k'|\mathcal{M}',\mathcal{P})\ \forall\mathcal{P} \implies p(k|\mathcal{M},\lambda) = p(k'|\mathcal{M}',\lambda)\ \forall\lambda$$

In words: operationally equivalent measurements are ontologically equivalent.

9.2 Preparation noncontextuality

We can notice that there are also operationally equivalent preparations. (In quantum theory, these are preparations represented by the same density operator.) Why not also impose that these operational equivalences imply ontological equivalence?

$$p(k|\mathcal{M},\mathcal{P}) = p(k|\mathcal{M},\mathcal{P}')\ \forall k,\mathcal{M} \implies p(\lambda|\mathcal{P}) = p(\lambda|\mathcal{P}')\ \forall\lambda$$

9.3 Outcome determinism

The assumption of preparation noncontextuality *implies* outcome determinism for certain measurements, thus answering the second question and then extending the principle involved gives an answer to the first question (which measurements are "projective") for free.

I will sketch the proof within quantum theory for the simple example of Pauli X and Z measurements on a qubit. The crucial property turns out to be that for each outcome of these measurements there exists a state that makes that outcome certain. For example, if we prepare $|0\rangle$ and then measure Pauli Z we get $+1$ with certainty:

$$p(+1|\mathcal{M}=Z,\mathcal{P}=|0\rangle) = 1$$

This means that for states $\Lambda_0 := \{\lambda : p(\lambda|\mathcal{P}=|0\rangle) > 0\}$ we must have $p(+1|\mathcal{M}=Z) = 1$. Similarly $|1\rangle$ makes the -1 outcome certain so for $\Lambda_1 := \{\lambda : p(\lambda|\mathcal{P}=|1\rangle) > 0\}$ we have $p(-1|\mathcal{M}=Z) = 1$. Overall then we have outcome determinism for the Z measurement for all the ontic states in $\Lambda_0 \cup \Lambda_1$.

For the X measurement we have preparations $|+\rangle$ and $|-\rangle$ that makes the outcome certain. Hence by a similar argument we have outcome determinism for the X measurement for $\Lambda_+ \cup \Lambda_-$ where $\Lambda_\pm = \{\lambda|p(\lambda|\mathcal{P}=|\pm\rangle) > 0\}$.

Now comes the application of preparation noncontextuality. We notice that tossing a fair coin and then preparing $|0\rangle$ or $|1\rangle$ is operationally equivalent to tossing a fair coin and then preparing $|+\rangle$ or $|-\rangle$:

$$\frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} = \frac{|+\rangle\langle +| + |-\rangle\langle -|}{2}$$

By preparation noncontextuality this operational equivalence implies ontological equivalence:

$$\frac{p(\lambda|\mathcal{P} = |0\rangle) + p(\lambda|\mathcal{P} = |1\rangle)}{2} = \frac{p(\lambda|\mathcal{P} = |0\rangle) + p(\lambda|\mathcal{P} = |1\rangle)}{2}$$

The LHS is nonzero on $\Lambda_0 \cup \Lambda_1$, where we know the Z measurement is outcome deterministic. The RHS is nonzero on $\Lambda_+ \cup \Lambda_-$, where we know X is outcome deterministic. Since the LHS equals the RHS, these are the same sets and so we have outcome determinism for both measurements on the same ontic states.

9.4 Noisy measurements

The results of the previous section can be generalised to show that if a measurement is close to being perfectly predictable then outcome determinism approximately holds on average [12].

If an operational equivalence of procedures fails to hold exactly, convexity can be used to find procedures for which it does hold exactly [13].

Together these answer the quantitative versions of both questions.

References

[1] S. Kochen and E. Specker, Indiana Univ. Math. J. **17**, 59 (1968).

[2] N. D. Mermin, Phys. Rev. Lett. **65**, 3373 (1990).

[3] J. Von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, 1955) first published as *Mathematische Grundlagen der Quantenmechanik* in 1932.

[4] A. Gleason, Indiana Univ. Math. J. **6**, 885 (1957).

[5] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).

[6] A. Cabello, S. Severini, and A. Winter, Phys. Rev. Lett. **112**, 040401 (2014), arXiv:1010.2163 .

[7] A. Acín, T. Fritz, A. Leverrier, and A. B. Sainz, Comm. Math. Phys. **334**, 533 (2015), arXiv:1212.4084 .

[8] S. Abramsky and A. Brandenburger, New J. Phys. **13**, 113036 (2011), arXiv:1102.0264 .

[9] R. W. Spekkens, Found. Phys. **44**, 1125 (2014), arXiv:1312.3667 .

[10] G. Chiribella and X. Yuan, "Measurement sharpness cuts nonlocality and contextuality in every physical theory," (2014), arXiv:1404.3348 .

[11] R. W. Spekkens, Phys. Rev. A **71**, 052108 (2005), arXiv:quant-ph/0406166 .

[12] R. Kunjwal and R. W. Spekkens, Phys. Rev. Lett. **115**, 110403 (2015), arXiv:1506.04150 .

[13] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, Nat. Commun. **7**, 11780 (2016), arXiv:1505.06244 .