Title: PSI 2016/2017 Foundations of Quantum Mechanics (Review) - Lecture 7

Date: Jan 11, 2017 11:30 AM

URL: http://pirsa.org/17010039

Abstract:

Quantum Foundations: Lecture 6

4) Ontological Models

Today's Lecture

4. Ontological Models
    3. Examples
    4. Excess Baggage Theorems
    5. Contextuality

Review of Last Lecture

- An ontological model of a prepare and measure experiment consists of:
  - A specification of an ontic state space Λ
  - For every preparation P, a probability distribution $Pr(\lambda|P)$ called the epistemic state.
  - For every measurement M, a probability distribution over the outcomes k that depends on λ $\quad Pr(k|M,\lambda)$

(Note: Thought of as a function of λ for fixed k, $Pr(k|M,\lambda)$ is often called the response function for M,k)

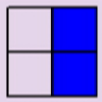
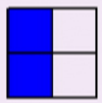
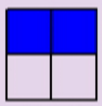
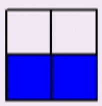
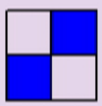
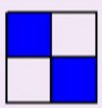
  - The model reproduces the operational predictions if

$$\mathrm{Prob}(k|P,M) = \int_\Lambda d\lambda \, Pr(k|M,\lambda) Pr(\lambda|P) \quad \left[= \mathrm{Tr}\left(E_k^M \rho_P\right) \text{ if there is a quantum model} \right]$$

4.3) Examples of Ontological Models

Spekkens' Toy Theory

$\Lambda = \{++, +-, -+, --\}$

Preparation P	Epistemic State $Pr(\lambda\|P)$
$+x$	
$-x$	
$+y$	
$-y$	
$+z$	
$-z$	

Measurement M	Outcome k	Response Function $Pr(k\|M,\lambda)$
X	$+1$	$\begin{array}{\|c\|c\|}\hline 0 & 1 \\ \hline 0 & 1 \\ \hline \end{array}$
X	-1	$\begin{array}{\|c\|c\|}\hline 1 & 0 \\ \hline 1 & 0 \\ \hline \end{array}$
Y	$+1$	$\begin{array}{\|c\|c\|}\hline 1 & 1 \\ \hline 0 & 0 \\ \hline \end{array}$
Y	-1	$\begin{array}{\|c\|c\|}\hline 0 & 0 \\ \hline 1 & 1 \\ \hline \end{array}$
Z	$+1$	$\begin{array}{\|c\|c\|}\hline 0 & 1 \\ \hline 1 & 0 \\ \hline \end{array}$
Z	-1	$\begin{array}{\|c\|c\|}\hline 1 & 0 \\ \hline 0 & 1 \\ \hline \end{array}$

The Beltrametti-Bugajski Model

- This is just an encoding of the orthodox interpretation in the ontological models framework.
- $\Lambda = \{$ set of projectors onto 1-d subspaces of $\mathcal{H} \}$
- When a pure state $|\psi\rangle$ is prepared

$$\Pr(\lambda|\psi) = \delta(|\lambda\rangle\langle\lambda| - |\psi\rangle\langle\psi|)$$

- When a POVM $M = \{E_k\}$ is measured

$$\Pr(k|M,\lambda) = \mathrm{Tr}(E_k|\lambda\rangle\langle\lambda|)$$

- By construction $\mathrm{Tr}(E_k|\psi\rangle\langle\psi|) = \int_\Lambda d\lambda \Pr(k|M,\lambda)\Pr(\lambda|\psi)$
- Note that different preparations of mixed states correspond to different epistemic states.

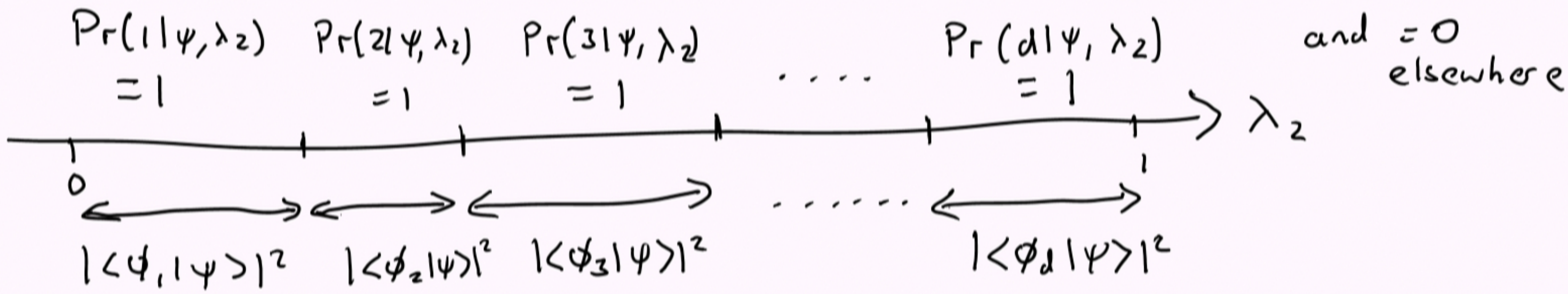
The Bell Model

- This was intended as an example to show that a deterministic hidden variable theory is always possible in principle
- $\Lambda = \underbrace{\{\text{set of projectors onto 1-d subspaces of } \mathcal{H}\}}_{\lambda_1} \times \underbrace{[0,1]}_{\lambda_2}$
- When pure state $|\psi\rangle$ is prepared

$$Pr(\lambda_1, \lambda_2 | \psi) = \delta(|\lambda_1\rangle\langle\lambda_1| - |\psi\rangle\langle\psi|)$$

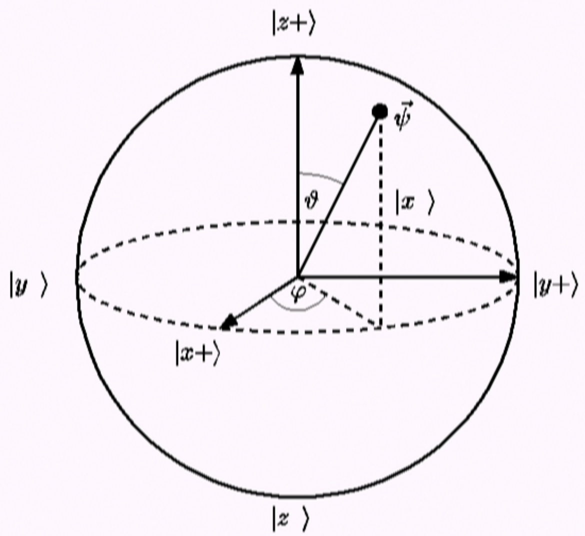
← uniform distribution over λ_2

- When an orthonormal basis $M = \{|\phi_k\rangle\langle\phi_k|\}_{k=1}^{d}$ is measured



The Kochen-Specker Model

- The Kochen-Specker model works for a qubit. We can use the Bloch sphere representation of state vectors.



Λ = the unit sphere

$$Pr(\vec{\lambda}\,|\,\vec{\psi}) = \frac{1}{\pi} H(\vec{\psi}\cdot\vec{\lambda})\,\vec{\psi}\cdot\vec{\lambda}$$

$$H(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

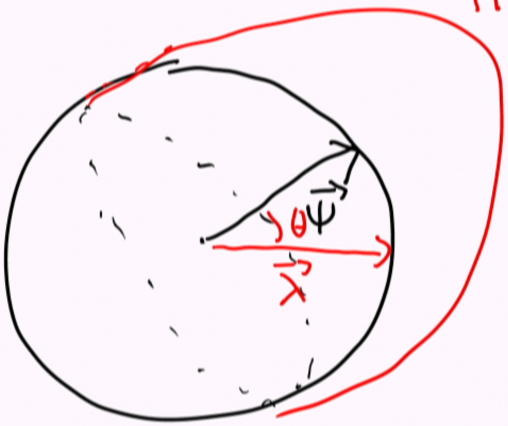
For an orthonormal basis measurement $M = \{|\phi\rangle\langle\phi|, |\phi^{\perp}\rangle\langle\phi^{\perp}|\}$

$$Pr(\vec{\phi}\,|\,\vec{\lambda}, M) = H(\vec{\phi}\cdot\vec{\lambda})$$

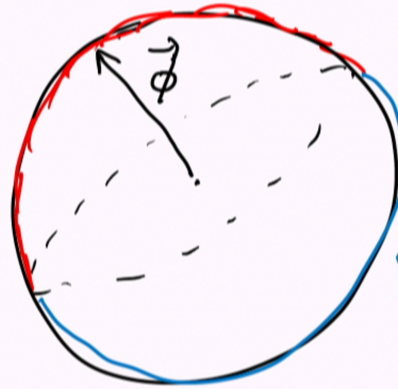
$$Pr(\vec{\phi}^{\perp}\,|\,\lambda, M) = 1 - H(\vec{\phi}\cdot\vec{\lambda})$$

Proving $|\langle\phi|\psi\rangle|^2 = \int_{\Lambda} d\lambda\, Pr(\vec{\phi}\,|\,\vec{\lambda}, M)\, Pr(\vec{\lambda}\,|\,\vec{\psi})$ is a hwk / tutorial exercise.

The Kochen-Specker Model



$Pr(\vec{\lambda}|\vec{\psi}) \propto \cos\theta$ in hemisphere of $\vec{\psi}$

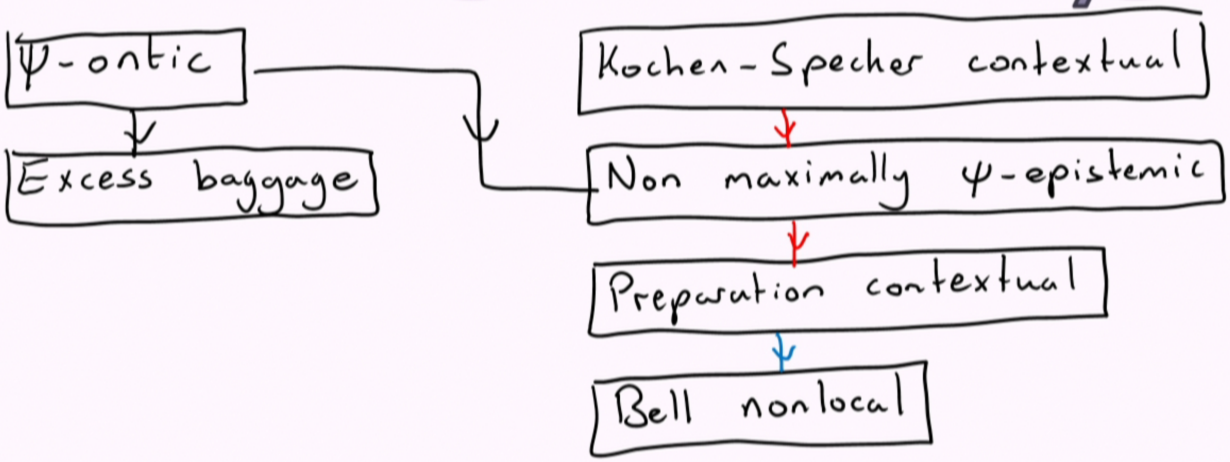


$Pr(\vec{\phi}|\lambda, M) = +1$

$Pr(\vec{\phi}|\lambda, M) = 0$

When computing integrals in KS model, don't forget that $d\vec{\lambda} = \sin\theta \, d\theta \, d\phi$ in spherical polar co-ordinates.

Hierarchy of Properties of Ontological Models of Quantum Theory



See M. Leifer, Quanta 3:67-155 (2014)

⊙ Some of the proofs use features of quantum theory that need not be true in arbitrary operational theories

→ = Uses self duality: $|\phi\rangle$ is both a state and a measurement outcome

→ = Uses steering: Any ensemble decomposition of a reduced density operator can be obtained by measuring the purifying system

4.4) Excess Baggage Theorems

Ontological Excess Baggage

- Lucien Hardy coined the term ontological excess baggage to refer to the fact that, even for a qubit Λ must be infinite.
- This is perhaps surprising because we can only reliably store and retrieve 1 bit of information in a qubit.
- Since Hardy's original proof, Montina has proved:
  - Λ must have the cardinality of the continuum.
  - Even if we allow the model to only approximately reproduce quantum theory, $|\Lambda| = O(e^d)$ where d is Hilbert space dimension
- Here we will prove Hardy's original result, see references in

  M. Leifer Quanta 3:67-155 (2014)

  D. Jennings and M. Leifer Contemp. Phys. 57:60-82 (2015)

  for references to later work.

A Useful Lemma

Lemma: Let P be a preparation of $|\psi\rangle$ and let M be a measurement in an orthonormal basis that includes $|\psi\rangle$

Let $\Lambda_\psi^P = \{\lambda \in \Lambda \,|\, \Pr(\lambda|P) > 0\}$, $\Gamma_\psi^M = \{\lambda \in \Lambda \,|\, \Pr(\psi|M,\lambda) > 0\}$

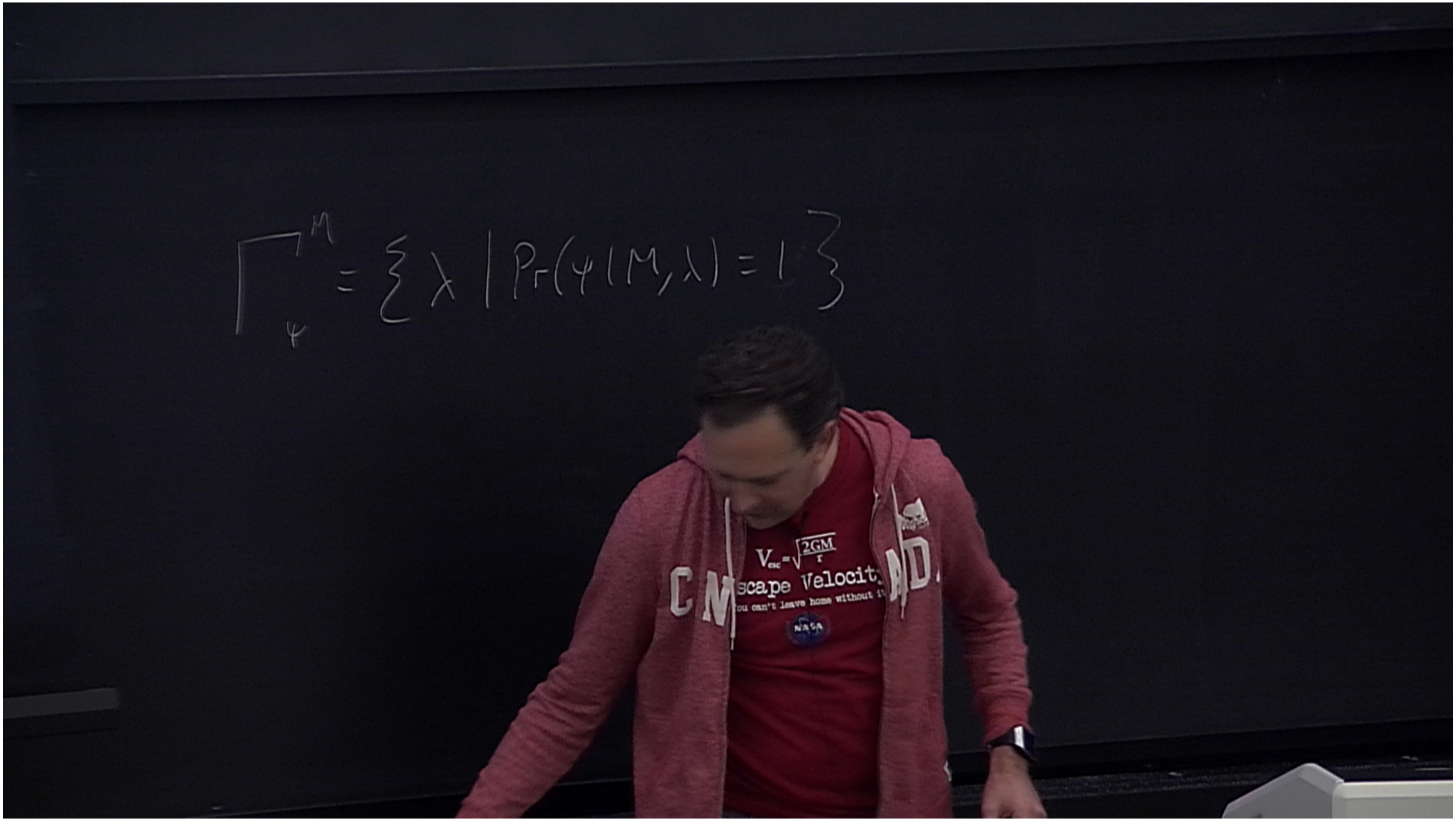
Then $\Lambda_\psi^P \subseteq \Gamma_\psi^M$ (up to measure-zero sets)

Proof:

$$1 = |\langle\psi|\psi\rangle|^2 = \int_\Lambda d\lambda \Pr(\psi|M,\lambda) \Pr(\lambda|P) = \int_{\Lambda_\psi^P} d\lambda \Pr(\psi|M,\lambda) \Pr(\lambda|P)$$

However, since $\int_{\Lambda_\psi^P} d\lambda \Pr(\lambda|P) = 1$ and $\Pr(\lambda|P) > 0$ on Λ_ψ^P

$\Pr(\psi|M,\lambda)$ must be 1 everywhere on Λ_ψ^P (up to measure zero sets)



$\Gamma_{\psi}^{M} = \{ \lambda \mid \Pr(\psi \mid M, \lambda) = 1 \}$

Hardy's Excess Baggage Theorem

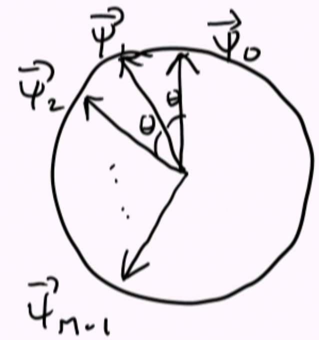
<u>Theorem</u>: Any ontological model that can reproduce the quantum predictions for orthonormal basis measurements on pure states in any Hilbert space dimension must have $|\Lambda| = \infty$.

<u>Proof</u>: Assume by contradiction that $|\Lambda| = N$ for some finite N.

Consider a 2d subspace spanned by $|0\rangle$ and $|1\rangle$ and the M states

$$|\psi_j\rangle = \cos\left(\frac{j\pi}{2M}\right)|0\rangle + \sin\left(\frac{j\pi}{2m}\right)|1\rangle \quad \text{for } j = 0, 1, 2, \ldots, M-1$$

We can choose M as large as we like



We have

$$|\langle\psi_k|\psi_j\rangle|^2 < 1 \quad \text{for all } j \neq k.$$

Hardy's Excess Baggage Theorem

Consider preparing the system in the state $|\psi_j\rangle$ and measuring it in a basis that includes $|\psi_k\rangle$ for $j \neq k$

Then $$\sum_{\lambda \in \Lambda} \Pr(\psi_k|\lambda) \Pr(\lambda|\psi_j) < 1$$

[Note: Naughty notation, so I'll use Λ_ψ for Λ_ψ^P also]

$\Rightarrow$ There must exist a $\lambda \in \Lambda_{\psi_j}$ s.t. $\Pr(\psi_k|\lambda) < 1$ otherwise sum would be 1

- Since $\Pr(\psi_k|\lambda) = 1$ everywhere on Λ_{ψ_k}, this means Λ_{ψ_j} and Λ_{ψ_k} must be different subsets of Λ.
- This applies to every pair $j \neq k$ so we must have M distinct subsets of Λ
- # distinct subsets of $\Lambda = 2^N \Rightarrow 2^N \geq M$ or $N \geq \log_2 M$
- But we can choose M as large as we like, so N is larger than any finite integer $\Rightarrow N = \infty$.

4.5) Contextuality

Leibniz Principle of the Identity of Indiscernables

- We follow an approach to contextuality that is due to Rob Spekkens
  – Phys. Rev. A 71, 052108 (2005).
- The basic philosophy is based on Leibniz Principle of the Identity of Indiscernables:
  - No two distinct things exactly resemble each other.
- This principle is arguably very successful in physics:
  - e.g. Principle of relativity, Einstein's equivalence principle.
- The principle can also be thought of as a no fine tuning argument.
  - e.g. suppose objects A and B have some distinct physical property, but there is absolutely no measurement we can do to tell A and B apart. Then, our measurements must only reveal coarse-grained information that is fine-tuned in just such a way so as not to reveal the difference.
- Not all apparent fine tunings are evil, but they do require explanation.

Preparation Contextuality

- Define an equivalence relation on preparations in an operational theory:

  $P \sim Q \quad \Leftrightarrow \quad \mathrm{Prob}(k|P,M) = \mathrm{Prob}(k|Q,M)$ for all measurement-outcome pairs (M,k).

- In particular, if $\rho_P = \rho_Q$ then $P \sim Q$.
- An ontological model is preparation noncontextual if,

$$P \sim Q \quad \Rightarrow \quad \Pr(\lambda|P) = \Pr(\lambda|Q).$$

- In words, whenever there is no observable distinction between two preparations, they are represented by the same epistemic state in the ontological model.
- A model that is not preparation noncontextual is called preparation contextual.

Mixing Preparations

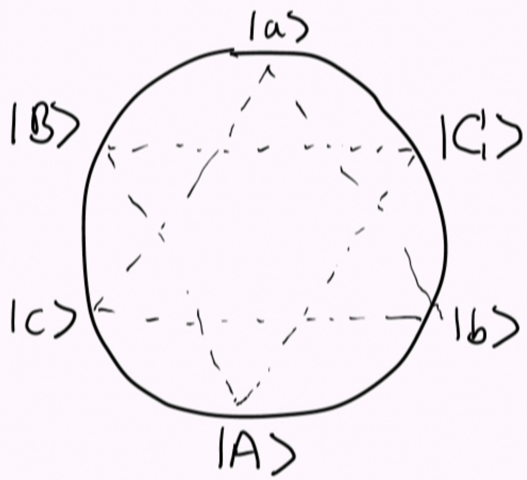
- If an operational theory contains preparations P and Q then we can construct a mixed preparation $pP + (1-p)Q$.
  - Physically this means, toss a coin with $p(\text{heads}) = p$, do P if it lands heads or Q if it lands tails, then forget the coin toss outcome.
- We will assume that the ontological model preserves mixtures:

$$\Pr(\lambda|pP + (1-p)Q) = p\Pr(\lambda|P) + (1-p)\Pr(\lambda|Q)$$

- This is actually an instance of preparation noncontextuality applied to the joint coin-system system. Conditioning on the outcome of the coin yields a preparation equivalent to P or Q.

Proof of Preparation Contextuality

- Consider the following 6 states on the equator of the Bloch sphere



We have $\langle a|A\rangle = \langle b|B\rangle = \langle c|C\rangle = 0$

This implies $\Lambda_a \cap \Lambda_A = \Lambda_b \cap \Lambda_B = \Lambda_c \cap \Lambda_C = \phi$

where $\Lambda_\psi = \{ \lambda \mid Pr(\lambda|\psi) > 0 \}$

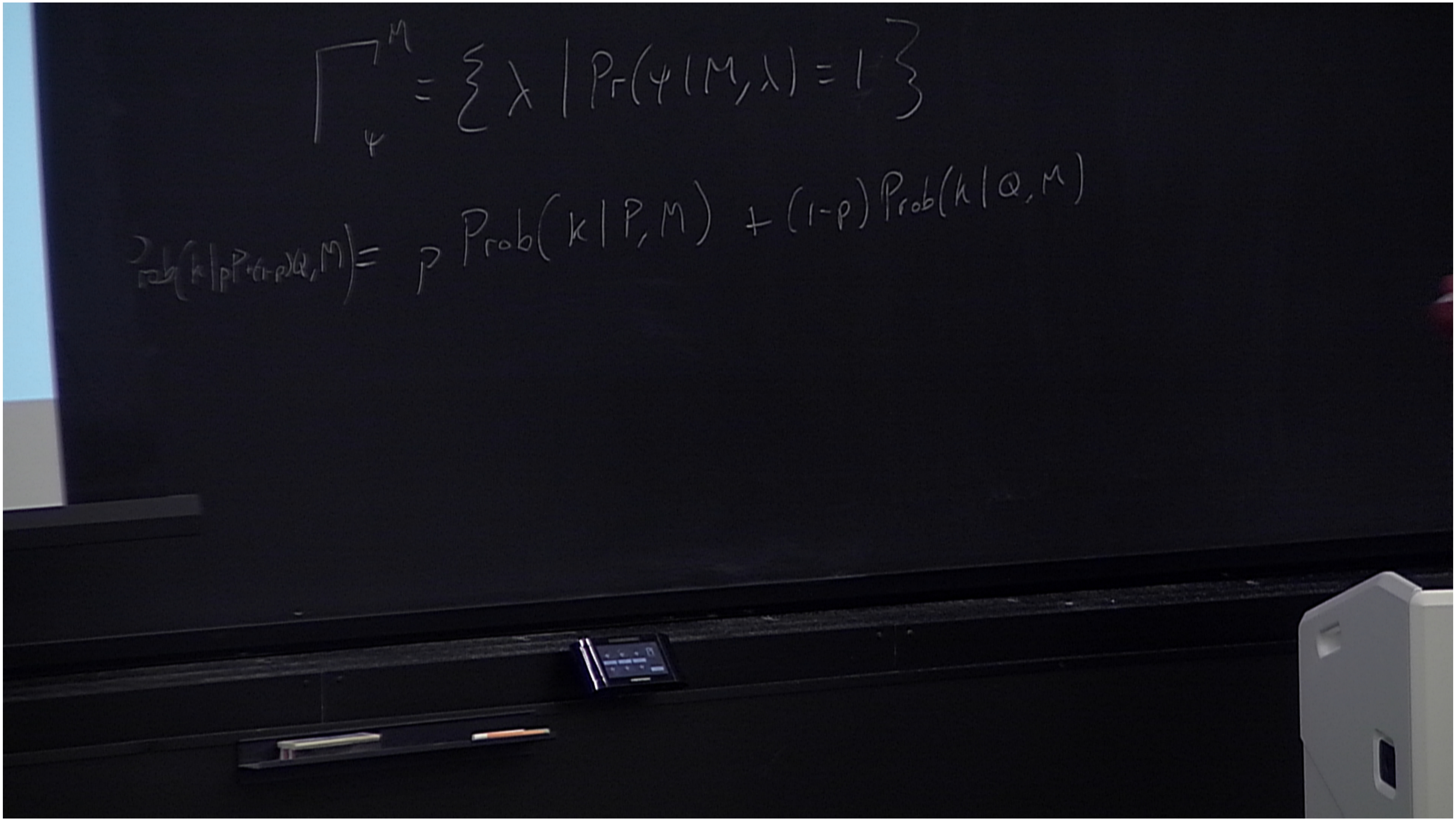
Why?

By lemma $\Lambda_\psi \subseteq \Gamma_\psi^M$ where $\Gamma_\psi^M = \{ \lambda \mid Pr(\psi|M,\lambda) = 1 \}$

But $Pr(a|M,\lambda) + Pr(A|M,\lambda) = 1$ for all $\lambda \in \Lambda$

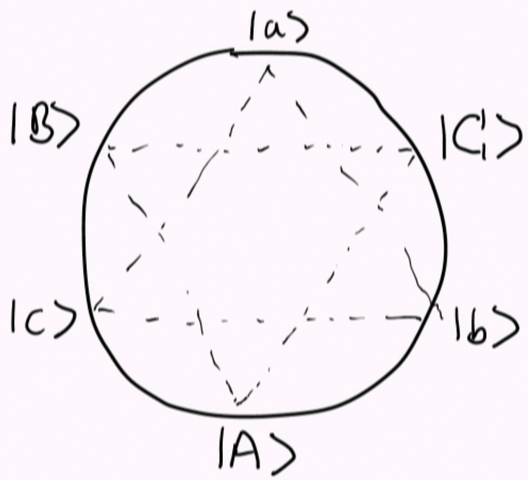
so $Pr(a|M,\lambda) = 1 \Rightarrow Pr(A|M,\lambda) = 0$ and vice versa

$\therefore \Gamma_a^M \cap \Gamma_A^M = \phi \Rightarrow \Lambda_a \cap \Lambda_A = \phi$


$\Gamma_\psi^M = \{\lambda \mid \Pr(\psi|M,\lambda)=1\}$
$\mathrm{Prob}(k|pP+(1-p)Q,M) = p\,\mathrm{Prob}(k|P,M) + (1-p)\,\mathrm{Prob}(k|Q,M)$

Proof of Preparation Contextuality

- Consider the following 6 states on the equator of the Bloch sphere



We have $\langle a|A\rangle = \langle b|B\rangle = \langle c|C\rangle = 0$

This implies $\Lambda_a \cap \Lambda_A = \Lambda_b \cap \Lambda_B = \Lambda_c \cap \Lambda_C = \phi$

where $\Lambda_\psi = \{\lambda \,|\, \Pr(\lambda|\psi) > 0\}$

Why?

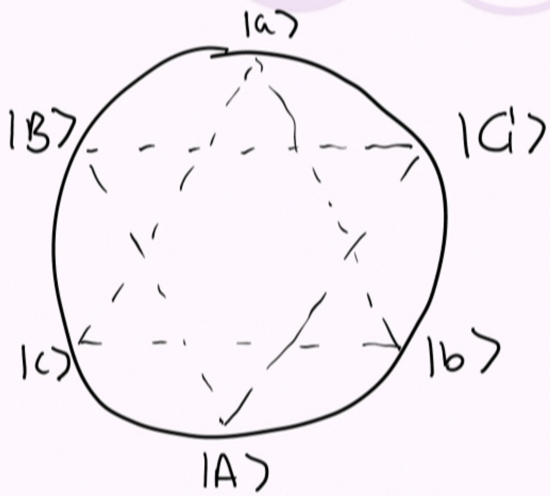
By lemma $\Lambda_\psi \subseteq \Gamma_\psi^M$ where $\Gamma_\psi^M = \{\lambda \,|\, \Pr(\psi|M,\lambda) = 1\}$

But $\Pr(a|M,\lambda) + \Pr(A|M,\lambda) = 1$ for all $\lambda \in \Lambda$

so $\Pr(a|M,\lambda) = 1 \Rightarrow \Pr(A|M,\lambda) = 0$ and vice versa

$\therefore \Gamma_a^M \cap \Gamma_A^M = \phi \Rightarrow \Lambda_a \cap \Lambda_A = \phi$

Proof of Preparation Contextuality



We also have:

$$
\begin{aligned}
\frac{I}{2} &= \frac{1}{2}\left(|a\rangle\langle a| + |A\rangle\langle A|\right) \\
&= \frac{1}{2}\left(|b\rangle\langle b| + |B\rangle\langle B|\right) \\
&= \frac{1}{2}\left(|c\rangle\langle c| + |C\rangle\langle C|\right) \\
&= \frac{1}{3}\left(|a\rangle\langle a| + |b\rangle\langle b| + |c\rangle\langle c|\right) \\
&= \frac{1}{3}\left(|A\rangle\langle A| + |B\rangle\langle B| + |C\rangle\langle C|\right)
\end{aligned}
$$

So by preparation noncontextuality:

$$
\begin{aligned}
\Pr(\lambda|\tfrac{I}{2}) &= \frac{1}{2}\left(\Pr(\lambda|a) + \Pr(\lambda|A)\right) \\
&= \frac{1}{2}\left(\Pr(\lambda|b) + \Pr(\lambda|B)\right) \\
&= \frac{1}{2}\left(\Pr(\lambda|c) + \Pr(\lambda|C)\right)
\end{aligned}
$$

$$
\begin{aligned}
\Pr(\lambda|\tfrac{I}{2}) &= \frac{1}{3}\left(\Pr(\lambda|a) + \Pr(\lambda|b) + \Pr(\lambda|c)\right) \\
&= \frac{1}{3}\left(\Pr(\lambda|A) + \Pr(\lambda|B) + \Pr(\lambda|C)\right)
\end{aligned}
$$

Proof of Preparation Contextuality

$$\Pr(\lambda|\tfrac{I}{2}) = \tfrac{1}{2}\left(\Pr(\lambda|a) + \Pr(\lambda|A)\right)$$
$$= \tfrac{1}{2}\left(\Pr(\lambda|b) + \Pr(\lambda|B)\right)$$
$$= \tfrac{1}{2}\left(\Pr(\lambda|c) + \Pr(\lambda|C)\right)$$

$$\Pr(\lambda|\tfrac{I}{2}) = \tfrac{1}{3}\left(\Pr(\lambda|a) + \Pr(\lambda|b) + \Pr(\lambda|c)\right)$$
$$= \tfrac{1}{3}\left(\Pr(\lambda|A) + \Pr(\lambda|B) + \Pr(\lambda|C)\right)$$

Now, any given λ can only be in at most one of Λ_a or Λ_A, Λ_b or Λ_B, Λ_c or Λ_C.

Let's choose a λ that is not in Λ_a, not in Λ_b, and not in Λ_C. Then

$$\Pr(\lambda|\tfrac{I}{2}) = \tfrac{1}{2}\Pr(\lambda|c) \qquad \Pr(\lambda|\tfrac{I}{2}) = \tfrac{1}{3}\Pr(\lambda|c)$$

$\Rightarrow 2\Pr(\lambda|\tfrac{I}{2}) = 3\Pr(\lambda|\tfrac{I}{2}) \Rightarrow \Pr(\lambda|\tfrac{I}{2}) = 0$ for this particular λ

We get a similar result for every choice of not in $\Lambda_{a/A}$, $\Lambda_{b/B}$, $\Lambda_{c/C}$

This exhausts $\Lambda \Rightarrow \Pr(\lambda|\tfrac{I}{2}) = 0$ everywhere, but this cannot be true for a probability distribution.

Measurement Contextuality

- Define an equivalence relation on measurement-outcome pairs in an operational theory:

  $(\mathrm{M},\mathrm{k})\sim(N,l) \quad \Leftrightarrow \quad \mathrm{Prob}(k|P,M)=\mathrm{Prob}(l|P,N)$ for all preparations P.
- In particular, if $E_k^M = E_l^N$ then $(\mathrm{M},\mathrm{k})\sim(M,l)$.
- An ontological model is measurement noncontextual if,

$$(M,k)\sim(\mathrm{N,l}) \quad \Rightarrow \quad \Pr(k|M,\lambda)=\Pr(l|N,\lambda).$$

- In words, whenever there is no observable distinction between two measurement-outcome pairs, they are represented by the same response function in the ontological model.
- A model that is not preparation noncontextual is called measurement contextual.

Kochen-Specker Contextuality

- Measurement noncontextual models exist:
  - e.g. Beltrametti-Bugajski: $\Pr(k|M,\lambda) = \mathrm{Tr}(\, E_k^M \,|\lambda\rangle\langle\lambda|)$.
- A Kochen-Specker (KS) noncontextual model is:
  - A model that only contains projective measurements.
  - Measurement noncontextual.
  - Outcome deterministic: $\Pr(\Pi|\lambda) = 0 \text{ or } 1$ for all λ.
- We will prove next time that:

$$\text{KS contextual} \;\Rightarrow\; \text{maximally } \psi\text{−epistemic} \;\Longrightarrow\; \text{preparation contextual}$$

  so KS contextuality is still worth proving.
- KS contextuality can only be proved in $d \geq 3$.
- By applying KS noncontextuality for projective measurements and measurement noncontextuality for POVMs, Spekkens obtained a proof in $d = 2$. We will focus on traditional KS proofs.

KS Contextuality and value assignments

- Due to the outcome determinism assumption, each λ determines a value function v_λ that assigns a value 0 or 1 to each projector.

$$v_\lambda(\Pi) = \Pr(\Pi|\lambda)$$

- Since probabilities must sum to 1, in each projective measurement $\{\Pi_k\}$, exactly one of the projectors must get value 1, the others getting value 0.
- We can also think of the value functions as assigning definite values to observables (self-adjoint operators) via

$$v(M) = \sum_j m_j \, v(\Pi_j)$$

KS Contextuality and value assignments

- Now, if two observables M and N commute then they have a joint eigendecomposition.

$$M = \sum_j m_j \Pi_j \qquad\qquad N = \sum_j n_j \Pi_j$$

- And we will have:

$$MN = \sum_j m_j n_j \Pi_j \qquad\qquad M + N = \sum_j (m_j + n_j)\Pi_j$$

- Since, in all of these decompositions, the same projector will get the value 1, whenever $[M, N] = 0$, the value functions will obey

$$v(MN) = v(M)v(N) \qquad\qquad v(M+N) = v(M) + v(N)$$

- If we define functions of operators by power series, this implies that whenever $M_1, M_2, \ldots$ all mutually commute then

$$v\big(f(M_1, M_2, \ldots)\big) = f(v(M_1), v(M_2), \ldots)$$

- So another way of defining KS noncontextuality is: there exists a value function that assigns eigenvalues to observables that obeys $v\big(f(M_1, M_2, \ldots)\big) = f(v(M_1), v(M_2), \ldots)$ for mutually commuting observables.

The Peres-Mermin Square

- Consider the following table of 9 two qubit observables:

			Row products
$\sigma_1 \otimes \sigma_1$	$\sigma_1 \otimes I$	$I \otimes \sigma_1$	I
$\sigma_3 \otimes \sigma_3$	$I \otimes \sigma_3$	$\sigma_3 \otimes I$	I
$\sigma_2 \otimes \sigma_2$	$\sigma_1 \otimes \sigma_3$	$\sigma_3 \otimes \sigma_1$	$-I$
I	I	I	Column products

- Each observable has eigenvalues ± 1, so receives values ± 1.
- Each row and column consists of mutually commuting observables.
- The column products are all $+I$, which has value $+1$, so there must be an even number of -1's in each column, so an even number in total
- However, one of the row products is $-I$, so there must be an odd number of -1's in that row, and an odd number in total $\Rightarrow$ contradiction.