

125c, Lecture Eleven: 5/8/17

Challenges for MWI:

- ① Branching (last time),
- ② Structure,
- ③ Probability.

Structure

Ordinarily we construct a quantum theory by starting w/ a classical model and quantizing it. E.g. a non-relativistic particle with phase space $\{\vec{x}, \vec{p}\}$, we make wave functions $\psi(\vec{x})$ and operators $\hat{x} = x$, $\hat{p} = i \frac{\partial}{\partial x}$. Square-integrable wave functions form a Hilbert space.

Of course we can make whatever operators we want, but somehow \hat{x} and \hat{p} seem preferred/natural.

But Nature (presumably) doesn't quantize ② anything; it's quantum right from the start. In MWI we have a Hilbert space and a Hamiltonian:

$$|\psi\rangle \in \mathcal{H}, \quad \hat{H}|\psi\rangle = i\partial_t |\psi\rangle.$$

\mathcal{H} is just an abstract vector space, with no preferred basis. In the absence of any preferred basis, \hat{H} is simply defined by its eigenvalues.

The "structure problem" in MWI is: how do we go from an abstract vector in a featureless Hilbert space to a rich phenomenological world of spatial locality, particles, fields, subsystems, etc.?

Two sub-problems:

① Factorization: how do we decompose $\mathcal{H} = \bigotimes_n \mathcal{H}_n$ in the best way?

② Preferred basis: given a factorization, what determines the basis $\{|a_n\rangle\}$ for individual factors?

Let's tackle the preferred-basis problem 3
first. (We've basically already solved it.)

Given e.g. $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$, of course we can use whatever basis we want for \mathcal{H}_S . But in practice we observe some but not others - Schrödinger's Cat is awake or asleep, not in a superposition. Why are these states preferred?

Of course we know the answer - decoherence.

The preferred basis consists of pointer states, robust under monitoring by the environment.

Therefore, it's the Hamiltonian that fixes the preferred basis. Given $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$, we write

$$\hat{H} = \hat{H}_S + \hat{H}_E + \hat{H}_I,$$

and pointer states satisfy

$$[|\phi_a\rangle\langle\phi_a|, \hat{H}_I] = 0.$$

So where did the decomposition $\mathcal{H}_S \otimes \mathcal{H}_E$ come from in the first place? It can also be thought of as a choice of basis. Given bases

$$\{|\omega_n\rangle\} \in \mathcal{H}, \quad \{|\phi_a\rangle\} \in \mathcal{H}_S, \quad \{|\eta_i\rangle\} \in \mathcal{H}_E,$$

what we really have is a vector-space isomorphism map

$$\Phi: \mathcal{H} \rightarrow \mathcal{H}_S \otimes \mathcal{H}_E$$

$$\begin{pmatrix} |\omega_1\rangle \\ |\omega_2\rangle \\ \vdots \end{pmatrix} \mapsto \begin{pmatrix} |\phi_1\rangle|\eta_1\rangle \\ |\phi_1\rangle|\eta_2\rangle \\ \vdots \end{pmatrix}$$

But of course this isomorphism is highly arbitrary. We can generate others by composing Φ with a unitary:

$$\Phi' = \Phi \circ \hat{U}.$$

This corresponds to a different way of decomposing:

$$\mathcal{H} = \mathcal{H}_S' \otimes \mathcal{H}_E'.$$

(4)

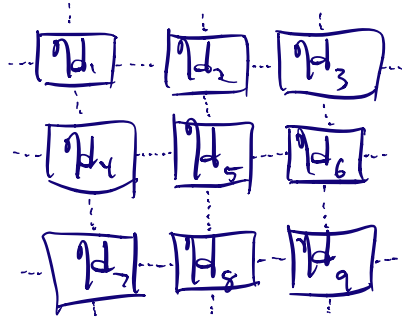
The "best" decomposition is determined, again, ⁽⁵⁾
 by the Hamiltonian. Two considerations:

① Locality. Consider a maximal decomposition
 into low-dimensional Hilbert spaces

$\mathcal{H} = \bigotimes_a \mathcal{H}_a$. "Locality" means that the
 interaction Hamiltonian will connect
 \mathcal{H}_i to \mathcal{H}_j if they are "nearby," and not
 if otherwise. Consider an expansion

$$\hat{H} = \sum_a h_a \hat{O}_a^{(\text{self})} + \sum_{ab} h_{ab} \hat{O}_{ab}^{(2)} + \sum_{abc} h_{abc} \hat{O}_{abc}^{(3)} \dots,$$

where each $\hat{O}_{a_1 \dots a_n}^{(n)}$
 acts nontrivially on n
 of the \mathcal{H}_a 's, and
 as the identity on others.



(Note that $\hat{O}_{a_1 \dots a_n}^{(n)}$ could itself be a sum
 over different operators acting on the
 relevant spaces. We're cheating by
 not keeping track of these parameters,
 but being careful yields the same result.)

⑥

Such a Hamiltonian is "k-local" if there are no terms involving more than k of the factors \mathcal{H}_α . If we think of \mathcal{H}_α 's as representing "degrees of freedom at some location in space," the real world appears local in this sense.

Key facts:

- Most Hamiltonians aren't k-local at all. Just count parameters. If $\dim \mathcal{H} = D$ and there are N \mathcal{H}_α 's of $\dim \mathcal{H}_\alpha = d$ each, then $D = d^N$.

With a bit of work (Cotler et al 2017), you can show that the number of free parameters in a k-local \hat{H} is much less than D , the number of eigenvalues of \hat{H} . Local Hamiltonians are special - a generic Hamiltonian can't be written in a k-local form.

- Even local Hamiltonians only look $\textcircled{1}$ local in the "right" factorizations. If we perform a random unitary on our tensor-product basis, then rewrite \hat{H} in the new basis, it generically won't look local at all.

Conclusion: given the happy fact that real-world physics appears local, there will be a preferred way of factorizing Hilbert space in which locality is manifest.

The other consideration relevant to the factorization problem is

$\textcircled{2}$ Robustness/weak-coupling.

We've noted that pointer states are robust under environmental monitoring.

Who says there should be any such states? In a generic decomposition

$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, we would expect all the states in \mathcal{H}_A to entangle with \mathcal{H}_B .

Schrödinger's Cat has a high-dimensional $\textcircled{8}$
Hilbert space - roughly $D = \exp(10^{24})$.

But we implicitly decompose

$$\mathcal{H}_{\text{cat}} = \mathcal{H}_{\text{macro}} \otimes \mathcal{H}_{\text{micro}},$$

where $\mathcal{H}_{\text{macro}}$ has a basis $\{ | \text{awake} \rangle, | \text{asleep} \rangle \}$
and $\mathcal{H}_{\text{micro}}$ is all the rest. Only $\mathcal{H}_{\text{macro}}$
becomes entangled with the environment.

That's clearly a factorization-dependent
statement. The "right" factorization
is one where entanglement proceeds
as slowly as possible. We can argue
(research in progress) that this corresponds
to decompositions where the interaction
terms \hat{H}_{int} are as small as possible.

Upshot: the origin of structure is an
interesting problem in MWI, but one
that is seemingly addressed by the
actual dynamics of the world.

The other big challenge for MWI - ⑨
the most serious one, really - is

Probability.

An obvious problem. In MWI we have

$$|\Psi(0)\rangle = (\alpha|\phi_1\rangle + \beta|\phi_2\rangle) |a_0\rangle |e_0\rangle$$

$$\rightsquigarrow |\Psi(T)\rangle = \alpha|\phi_1\rangle |a_1\rangle |e_1\rangle + \beta|\phi_2\rangle |a_2\rangle |e_2\rangle.$$

It's reasonable to ask: what do α and β have to do with anything? They don't seem to be related to the probability of anything happening. $|a_0\rangle$ evolves into both $|a_1\rangle$ and $|a_2\rangle$, with probability 1! MWI is a completely deterministic theory.

The problem is not really why $p(\alpha) = |\alpha|^2$ rather than some other function of α . Gleason's Theorem says basically that $p = |\alpha|^2$ is the unique probability measure satisfying some simple axioms.

The problem is why there are probabilities at all.

(10)

One (good) answer: **Self-locating Uncertainty**.

Consider Everettian evolution with an explicit observer (person) included:

$$|\Psi(0)\rangle = (\alpha|\phi_1\rangle + \beta|\phi_2\rangle) |a_0\rangle |e_0\rangle |p_0\rangle \quad \text{"ready"}$$

$$|\Psi(1)\rangle = (\alpha|\phi_1\rangle |a_1\rangle + \beta|\phi_2\rangle |a_2\rangle) |e_0\rangle |p_0\rangle \quad \text{"measurement"}$$

$$|\Psi(2)\rangle = (\alpha|\phi_1\rangle |a_1\rangle |e_1\rangle + \beta|\phi_2\rangle |a_2\rangle |e_2\rangle) |p_0\rangle \quad \text{"decoherence"}$$

$$|\Psi(3)\rangle = \alpha|\phi_1\rangle |a_1\rangle |e_1\rangle |p_1\rangle + \beta|\phi_2\rangle |a_2\rangle |e_2\rangle |p_2\rangle \quad \text{"observation"}$$

Decoherence is fast - much faster than a person can read out the apparatus.

Branching happens when the apparatus becomes entangled with the environment.

This stage can be written equivalently as

$$|\Psi(2)\rangle = \alpha|\phi_1\rangle |a_1\rangle |e_1\rangle |p_0\rangle + \beta|\phi_2\rangle |a_2\rangle |e_2\rangle |p_0\rangle.$$

Two branches, two observers, but both in identical states.

"Self-locating uncertainty" is just the fact that neither observer knows which branch they are on. There's something they don't know, even if they know the wave function of the universe exactly. (11)

Can they sensibly assign a probability to being in either branch?

More carefully: think of probability like a Bayesian - "subjective credence [degree of belief] that something is true" - rather than like a frequentist - "the fraction of times something happens in the limit of infinite trials." Is there a "right" way to assign such credences?

If there is a uniquely rational assignment of credences, Gleason's Theorem implies we expect it's the Born Rule. We'll give two arguments. (Maybe one good one would be sufficient?)

Assigning Credences to Branches

(12)

Strategy #1: "Epistemic Separability Principle"
[Sebens & Carroll]

Imagine we have a system, a person, an environment, and some extra part of the universe $\{|x_n\rangle\} \in \mathcal{H}_X$.

$$|\Psi_A\rangle = \alpha |d_1\rangle |x_1\rangle |e_1\rangle |p_0\rangle + \beta |d_2\rangle |x_2\rangle |e_2\rangle |p_0\rangle.$$

Person is the same on both branches, but everything else can be different.

ESP: whatever credence we assign to being on branch 1 or branch 2 should not depend on the state of \mathcal{H}_X .

I.e. imagine a different state, similar to $|\Psi_A\rangle$ but with different $|x_i\rangle$'s:

$$|\Psi_B\rangle = \alpha |d_1\rangle |x_3\rangle |e_1\rangle |p_0\rangle + \beta |d_2\rangle |x_4\rangle |e_2\rangle |p_0\rangle.$$

ESP doesn't (immediately) say anything about what credence our observers should associate with either branch. It just says they should be the same in $|\Psi_A\rangle$ and $|\Psi_B\rangle$.

$$P(k_1|A) = P(k_1|B),$$
$$P(k_2|A) = P(k_2|B).$$

Consider an explicit example, where all we do is to switch $|x_1\rangle \leftrightarrow |x_2\rangle$, and also set $\alpha = \beta = 1/\sqrt{2}$. (We had better get this right.)

$$|\Psi_A\rangle = \frac{1}{\sqrt{2}} |k_1\rangle |x_1\rangle |e_1\rangle |p_0\rangle + \frac{1}{\sqrt{2}} |k_2\rangle |x_2\rangle |e_2\rangle |p_0\rangle$$

$$|\Psi_B\rangle = \frac{1}{\sqrt{2}} |k_1\rangle |x_2\rangle |e_1\rangle |p_0\rangle + \frac{1}{\sqrt{2}} |k_2\rangle |x_1\rangle |e_2\rangle |p_0\rangle.$$

First, treat $|x_i\rangle$'s as part of the environment and ignore them. The remaining states (reduced density matrices) look the same in A & B.

Therefore e.g.

(14)

$$p(d_1|A) = p(d_1|B).$$

Now the trick: we could just as well imagine treating $|d_i\rangle$'s as part of the environment, and measuring the $|x_i\rangle$'s. Again, the reduced density matrices look the same, so we should attach equal credences in cases A & B. E.g.

$$p(x_2|A) = p(x_2|B).$$

But note that " $x_2|B$ " is the same physical branch as " $d_1|B$ " was. So

$$p(d_1|B) = p(x_2|B)$$

and
$$p(d_2|A) = p(x_2|A).$$

Putting them all together:

$$p(d_1|A) = p(d_2|A) = 1/2 = (1/\sqrt{2})^2.$$

As we would expect.