10. The Dynamics by Itself

1. Many Worlds
2. Bare Theory
3. Many Minds

- Consider composite system of human observer h, *Color* measuring device m, and electron e.
- *Suppose*: Pre-measurement state is $|ready\rangle_h|ready\rangle_m|ready\rangle_e$.
- *Then*: Schrödinger dynamics entails post-measurement state will be:

$$\sqrt{1/2}\,|believes\ e\ black\rangle_h|\text{"}black\text{"}\rangle_m|black\rangle_e + \sqrt{1/2}\,|believes\ e\ white\rangle_h|\text{"}white\text{"}\rangle_m|white\rangle_e$$

According to Option (A1), this is a state in which:

- e has no definite color.
- m has no definite reading.
- h has no definite belief about measurement outcome.

But: *According to our experience, measurements are supposed to have unique outcomes!*

10. The Dynamics by Itself

- Consider composite system of human observer *h*, *Color* measuring device *m*, and electron *e*.
- *Suppose*: Pre-measurement state is $|ready\rangle_h|ready\rangle_m|ready\rangle_e$.
- *Then*: Schrödinger dynamics entails post-measurement state will be:

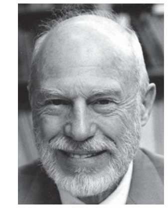
$$\sqrt{1/2}\,|believes\ e\ black\rangle_h|\text{"}black\text{"}\rangle_m|black\rangle_e + \sqrt{1/2}\,|believes\ e\ white\rangle_h|\text{"}white\text{"}\rangle_m|white\rangle_e$$

- *The Measurement Problem*: How to reconcile the *Schrödinger dynamics* with the *Projection Postulate*; *i.e.*, how to reconcile superpositions with our experience that *measurements have unique outcomes*.
- *GRW Solution*: Keep Projection Postulate and modify Schrödinger dynamics so that superpositions will *not* occur after measurements.
- *Dynamics-By-Itself Solutions*: Keep Schrödinger dynamics and give up Projection Postulate! Attempt to explain how measurements do not *really* have unique outcomes (even though we think they do).

1. The Many-Worlds (MW) Interpretation

Everett (1957), DeWitt (1970)

Hugh Everett III
(1930-1982)



Bryce DeWitt
(1923-2004)

MW Claims:

(A) States evolve only *via* Schrödinger dynamics (no Projection Postulate).

(B) Each term in a superposition represents a state in a *distinct world*.

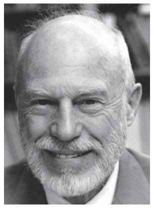
$$\sqrt{1/2}\,\underbrace{|believes\ e\ black\rangle_h|\text{"}black\text{"}\rangle_m|black\rangle_e}_{\text{state of } h\text{-}m\text{-}e \text{ system in World } A} + \sqrt{1/2}\,\underbrace{|believes\ e\ white\rangle_h|\text{"}white\text{"}\rangle_m|white\rangle_e}_{\text{state of } h\text{-}m\text{-}e \text{ system in World } B}$$

According to the EE Rule, in both Worlds A and B:

- e has a definite value of *Color*.
- m has a definite reading.
- h has a definite belief about the measurement outcome.

1. The Many-Worlds (MW) Interpretation Everett (1957), DeWitt (1970)

Hugh Everett III
(1930-1982)



Bryce DeWitt
(1923-2004)

MW Claims:

(A) States evolve only *via* Schrödinger dynamics (no Projection Postulate).

(B) Each term in a superposition represents a state in a *distinct world*.

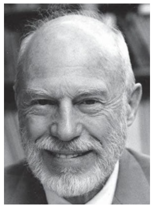
$$\underbrace{\sqrt{1/2}\,|believes\ e\ black\rangle_h|\text{"}black\text{"}\rangle_m|black\rangle_e}_{\text{state of } h\text{-}m\text{-}e \text{ system in World } A} + \underbrace{\sqrt{1/2}\,|believes\ e\ white\rangle_h|\text{"}white\text{"}\rangle_m|white\rangle_e}_{\text{state of } h\text{-}m\text{-}e \text{ system in World } B}$$

Consequences:

- Any given measurement does not have one unique outcome!
  - *When a measurement occurs, all its possible outcomes are realized, one per world.*
- Each time a measurement occurs, the world splits into as many worlds as there are possible measurement outcomes.
- There is *no interaction between worlds*.
  - *If we think measurements have unique outcomes, then we don't experience splits.*

1. The Many-Worlds (MW) Interpretation Everett (1957), DeWitt (1970)

Hugh Everett III
(1930-1982)



Bryce DeWitt
(1923-2004)

MW Claims:

(A) States evolve only *via* Schrödinger dynamics (no Projection Postulate).

(B) Each term in a superposition represents a state in a *distinct world*.

$$\sqrt{1/2}\,\underbrace{|\textit{believes e black}\rangle_h|\text{"}\textit{black}\text{"}\rangle_m|\textit{black}\rangle_e}_{\text{state of } h\text{-}m\text{-}e \text{ system in World } A} + \sqrt{1/2}\,\underbrace{|\textit{believes e white}\rangle_h|\text{"}\textit{white}\text{"}\rangle_m|\textit{white}\rangle_e}_{\text{state of } h\text{-}m\text{-}e \text{ system in World } B}$$

One way to think of this:

$$|\textit{universe}\rangle \xrightarrow[\text{Schrödinger evolution}]{} \sqrt{1/2}\,|\textit{universe}\rangle_A + \sqrt{1/2}\,|\textit{universe}\rangle_B$$

universal state vector at any given instant

universal state vectors for Worlds A and B after Color measurement

- *In general*: *Any* interaction between two or more physical systems may result in a splitting of worlds (since any interaction that is governed by the Schrödinger dynamics may result in a superposition.)

Problems

1. The Preferred Basis Problem

- A given superposition can always be rewritten in a different basis.

$$\overbrace{\sqrt{1/2}\,|\mathit{believes\ e\ black}\rangle_h|\text{"}\mathit{black}\text{"}\rangle_m|\mathit{black}\rangle_e}^{\textit{World A with black electron}} + \overbrace{\sqrt{1/2}\,|\mathit{believes\ e\ white}\rangle_h|\text{"}\mathit{white}\text{"}\rangle_m|\mathit{white}\rangle_e}^{\textit{World B with white electron}}$$

$$= \sqrt{1/2}\,\underbrace{|Q+\rangle_{h\&m}|\mathit{hard}\rangle_e}_{\textit{World C with hard electron}} + \sqrt{1/2}\,\underbrace{|Q-\rangle_{h\&m}|\mathit{soft}\rangle_e}_{\textit{World D with soft electron}}$$

$$|Q+\rangle_{h\&m} = \sqrt{1/2}\,|\mathit{believes\ e\ black}\rangle_h|\text{"}\mathit{black}\text{"}\rangle_m + \sqrt{1/2}\,|\mathit{believes\ e\ white}\rangle_h|\text{"}\mathit{white}\text{"}\rangle_m$$

$$|Q-\rangle_{h\&m} = \sqrt{1/2}\,|\mathit{believes\ e\ black}\rangle_h|\text{"}\mathit{black}\text{"}\rangle_m - \sqrt{1/2}\,|\mathit{believes\ e\ white}\rangle_h|\text{"}\mathit{white}\text{"}\rangle_m$$

- If we initially had a *hard* electron and we let it interact with a *Color* measuring device, what does *MW* say about how the world splits?
- Does it split into Worlds *A* and *B*, or does it split into Worlds *C* and *D*?

Task: *Find and justify a fundamental basis in terms of which all superpositions should be expanded.*

2. The Problem of Probabilities

$$\sqrt{½}\,|believes\ e\ black\rangle_h|\text{"}black\text{"}\rangle_m|black\rangle_e + \sqrt{½}\,|believes\ e\ white\rangle_h|\text{"}white\text{"}\rangle_m|white\rangle_e$$

- *Born Rule says*: When *h* measures the *Color* of *e*, there is a probability of ½ that the outcome will be *black*, and a probability of ½ that the outcome will be *white*.
- *MW says*: When *h* measures the *color* of *e*, the world splits into a world in which the outcome is *black* with *certainty*, and a world in which the outcome is *white* with *certainty*!

- *Born Rule says*: "Each outcome has a distinct probability of occurring."
- *MW says*: "All outcomes occur."

Where did the probabilities go in MW?

Possible Responses

(i) *MW probabilities are ontic*: They are defined over possible worlds.

- *So*: When a *Color* measurement is conducted, the world splits with probability ½ into World *A*, and probability ½ into World *B*.

- *But*: Need to specify a *dynamical law* that tells us how a given world evolves over time indeterministically into others.

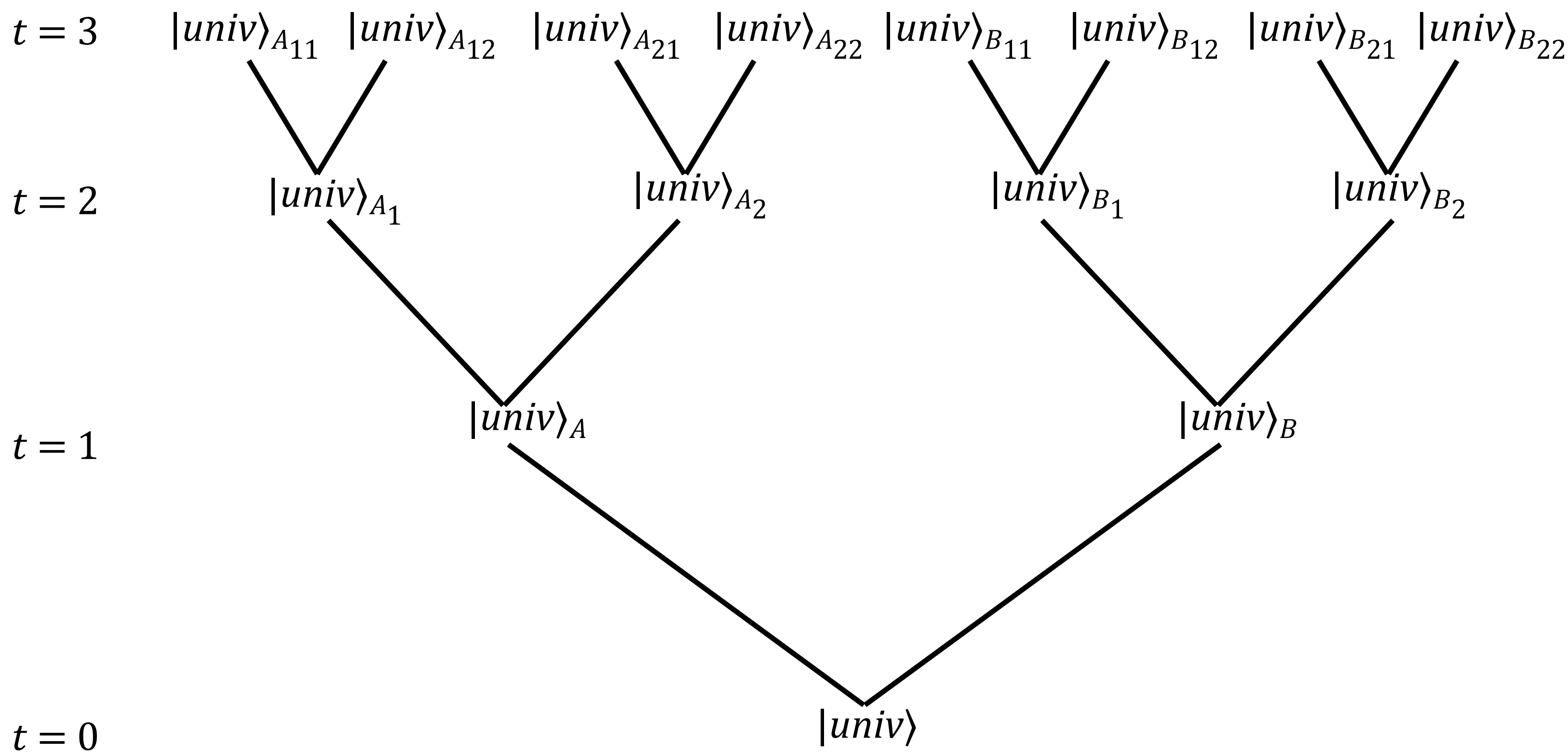
(ii) *MW probabilities are epistemic*: They reflect the state of knowledge of the human observer in the act of measurement.

- *So*: When a *Color* measurement is conducted, *h* doesn't know which world (*A* or *B*) she will end up in; she only knows the probability of which world she will end up in.

- *But*: Aren't there *two h*'s after the measurement, one in each world with certainty?
- *And*: Seems to fall back on distinction between *measurements* (interactions involving human observers) and other types of interactions, which is what *MW* rejects.

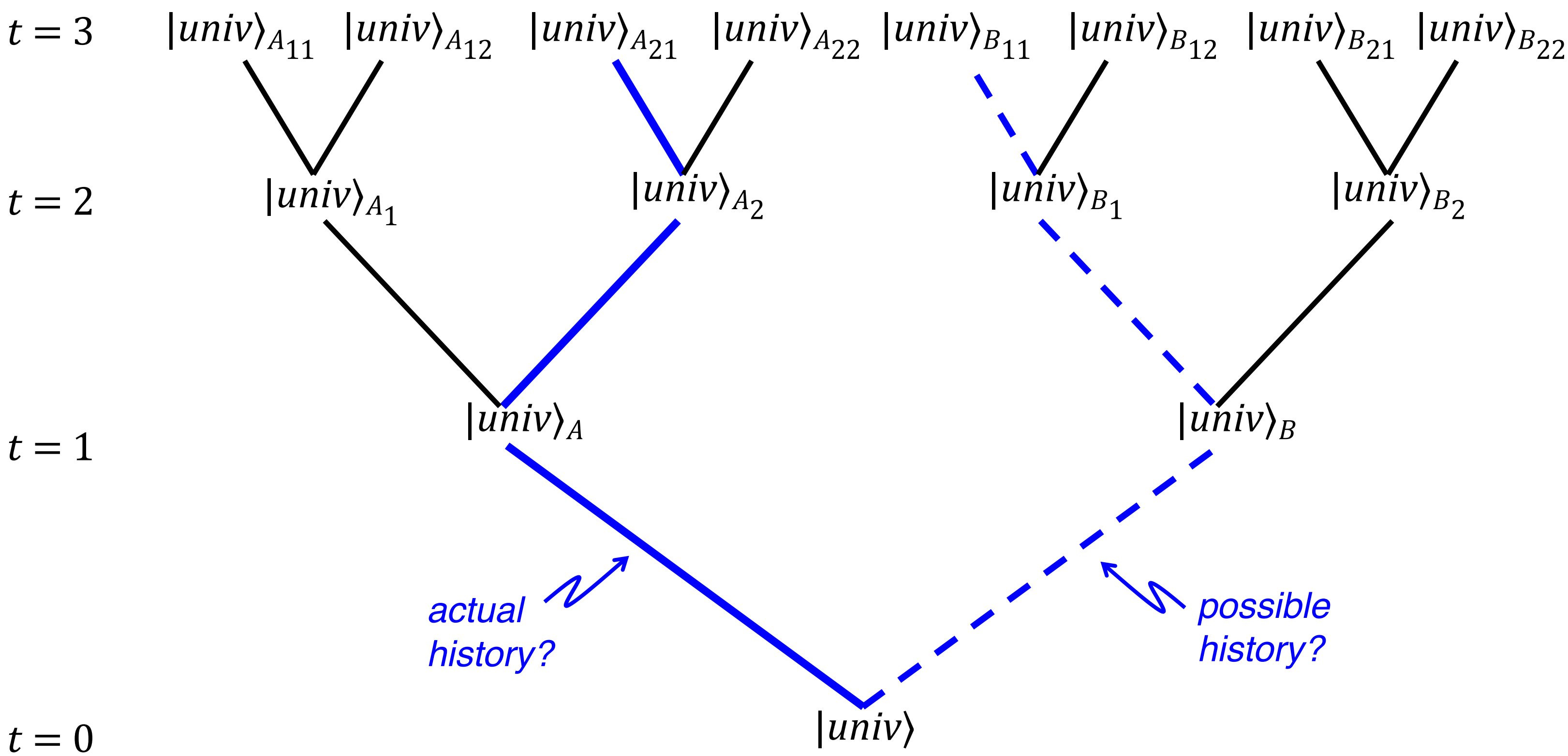
(1) How do we justify introducing probabilities (ontic or epistemic) into MW?

(2) How do we justify introducing the *correct* QM probabilities into MW?



(1) How do we justify introducing probabilities (ontic or epistemic) into MW?

(2) How do we justify introducing the *correct* QM probabilities into MW?



(1) Need to be able to pick out *possible histories* of worlds, and be then be able to distinguish them from the *actual history*.

(2) Need to be able to explain why this Rule assigns the correct QM probabilities to worlds.

3. The Problem of Conservation Laws

- When the universe splits, aren't conservation laws violated?

$$\left(\begin{array}{c}\textit{universe}\\\textit{splits}\end{array}\right) \Rightarrow \left(\begin{array}{c}\textit{number of physical}\\\textit{objects increases}\end{array}\right) \Rightarrow \left(\begin{array}{c}\textit{violation of conservation}\\\textit{of mass/energy?}\end{array}\right)$$

Possible Response:

- A world includes *spacetime* as well as physical objects.
- *And*: Worlds split "outside" of spacetime: each new world contains its own spacetime and its own physical objects.
- *So*: No violations of mass/energy conservation.

> - *But*: Does it make sense to say splits occur outside of time?
> - Don't we want to say something like: "At time t_1, there is one world; and at time $t_2 > t_1$, after a *Color* measurement, there are two worlds."

2. The Bare Theory Albert (1992)

David Albert

- *MW says*: Keep Schrödinger dynamics and give up Projection Postulate.
  - *Attempt to explain how measurements do not really have unique outcomes (even though we think they do).*
- *Bare Theory says*: The same thing, but differs on the explanation of why measurements don't really have unique outcomes.
  - *Attempts to offer an explanation without all the "world-talk".*

Bare Theory Claims:

(A) States evolve only *via* Schrödinger dynamics (no Projection Postulate).

(B) We are mistaken in thinking measurements have unique outcomes. (We never have definite beliefs about measurement outcomes; at best, we have "effective knowledge" about outcomes.)

- *Motivation*: What would it feel like to be in a superposition?

  Suppose h conducts a *Color* measurement on e with result $(*)$:

$$\sqrt{½}\,|believes\ e\ black\rangle_h|"black"\rangle_m|black\rangle_e + \sqrt{½}\,|believes\ e\ white\rangle_h|"white"\rangle_m|white\rangle_e$$

- *Ask h*: "Do you have any definite belief about the value of the *Color* of e?"

First Note:

- Suppose $\mathcal{O}$ is a linear operator representing some property and let $|A\rangle$ and $|B\rangle$ be eigenvectors of $\mathcal{O}$ with the same eigenvalue λ:

  $\mathcal{O}|A\rangle = \lambda|A\rangle, \quad \mathcal{O}|B\rangle = \lambda|B\rangle$.
- Then any linear superposition $\alpha|A\rangle + \beta|B\rangle$ of these eigenvectors will also be an eigenvector of $\mathcal{O}$ with eigenvalue λ (where α and β are numbers).

Proof:

$$\begin{aligned}\mathcal{O}(\alpha|A\rangle + \beta|B\rangle) &= \alpha(\mathcal{O}|A\rangle) + \beta(\mathcal{O}|B\rangle)\\ &= \alpha\lambda|A\rangle + \beta\lambda|B\rangle\\ &= \lambda(\alpha|A\rangle + \beta|B\rangle)\end{aligned}$$

- *Motivation*: What would it feel like to be in a superposition?

  Suppose h conducts a *Color* measurement on e with result $(*)$:

$$\sqrt{1/2}\,|\textit{believes e black}\rangle_h|\textit{"black"}\rangle_m|\textit{black}\rangle_e + \sqrt{1/2}\,|\textit{believes e white}\rangle_h|\textit{"white"}\rangle_m|\textit{white}\rangle_e$$

- *Ask h*: "Do you have any definite belief about the value of the *Color* of e?"

Now Note:

- If after measurement, the h-m-e state is given by the first term of $(*)$, then h will respond to the question with "*Yes*".
- If the h-m-e state is given by the second term of $(*)$, then h will respond with "*Yes*".

- Think of the question as a property of the h-m-e system and "*Yes*" as a value of this property.
- Since both terms in the superposition are states that have the value "*Yes*" of this property, the superposition itself is a state which has the value "*Yes*" of this property!

- *Motivation*: What would it feel like to be in a superposition?

  Suppose h conducts a *Color* measurement on e with result $(*)$:

$$\sqrt{1/2}\,|\textit{believes e black}\rangle_h|\textit{"black"}\rangle_m|\textit{black}\rangle_e + \sqrt{1/2}\,|\textit{believes e white}\rangle_h|\textit{"white"}\rangle_m|\textit{white}\rangle_e$$

- *Ask h*: "Do you have any definite belief about the value of the *Color* of e?"

So:
- By the *Eigenvector/Eigenvalue Rule*, when $(*)$ obtains, h doesn't have a definite belief about the *Color* of e.
- *But*: h "effectively knows" what the *Color* of e is: h will answer "*Yes*" if asked if she knows what the *Color* of e is.

- *Consequence*: According to the Bare Theory, we are *always* mistaken about the values of the properties of physical systems.
- The *only* beliefs that we are never mistaken about are beliefs about whether or not some definite measurement result was observed.

Problems with the Bare Theory

(1) Conflicts with Common Perceptions of Introspection.

- *Claim*: Our beliefs may be wrong, but we are certain that we hold them.

The Bare Theory denies this claim!

- Suppose you are *h*. Ask yourself: "Did I just see a definite *Color* result for *e*?"
- According to the Bare Theory, you will answer "*Yes*", but this is mistaken: You really do not have a definite belief about what the *Color* of *e* is because *your belief state is in a superposition*.

(2) Self-defeating?

- According to the Bare Theory, any belief we might have for evidence for it (or for quantum mechanics in general) would be mistaken.

3. The Many Minds (MM) Interpretation Albert & Loewer (1988)


David Albert


Barry Loewer

- *Bare Theory*: Tries to tell a story about how belief states in superpositions can still be said to have "effective" collapses, even if they really don't.
- *Many Minds*: Distinguishes between physical states and mental states and says, physical states can be in superpositions, but mental states never are.

> *MM Claims*:
>
> (A) *Physical states* evolve only *via* Schrödinger dynamics (no Projection Postulate).
>
> (B) *Mental states* ("minds") evolve *via* an indeterministic dynamics in such a way that they are never in superpositions.

- *Motivation*: Suppose h conducts a *Color* measurement on e and ends up in standard state:

$$\sqrt{½}\,|\mathit{believes\ e\ black}\rangle_h|\text{"}\mathit{black}\text{"}\rangle_m|\mathit{black}\rangle_e + \sqrt{½}\,|\mathit{believes\ e\ white}\rangle_h|\text{"}\mathit{white}\text{"}\rangle_m|\mathit{white}\rangle_e$$

MM says:

- This represents a *physical* state. In particular, the h-states are physical brain states of h.
- These *brain states* of h have corresponding *mental states* (which aren't represented in the standard state).
- The standard state represents the physical state in which h is in the mental state associated with the brain state $|\mathit{believes\ e\ black}\rangle_h$ with probability ½, and h is in the mental state associated with the brain state $|\mathit{believes\ e\ white}\rangle_h$ with probability ½.

- *Motivation*: Suppose h conducts a *Color* measurement on e and ends up in standard state:

$$\sqrt{1/2}\,|\textit{believes e black}\rangle_h|\textit{"black"}\rangle_m|\textit{black}\rangle_e + \sqrt{1/2}\,|\textit{believes e white}\rangle_h|\textit{"white"}\rangle_m|\textit{white}\rangle_e$$

Why this is supposed to help:

- Suppose h is in the physical state represented by the standard state and has a definite belief about what the *Color* of e is.
- According to the *Eigenvector/Eigenvalue Rule*, this is a *false* belief.
- But, according to MM, h's mental state evolves in a way that is consistent with supposing h's belief were true. For instance, if she thinks e is *black*, then her mental state evolves to the mental state corresponding to the brain state $|\textit{believes e black}\rangle_h$.

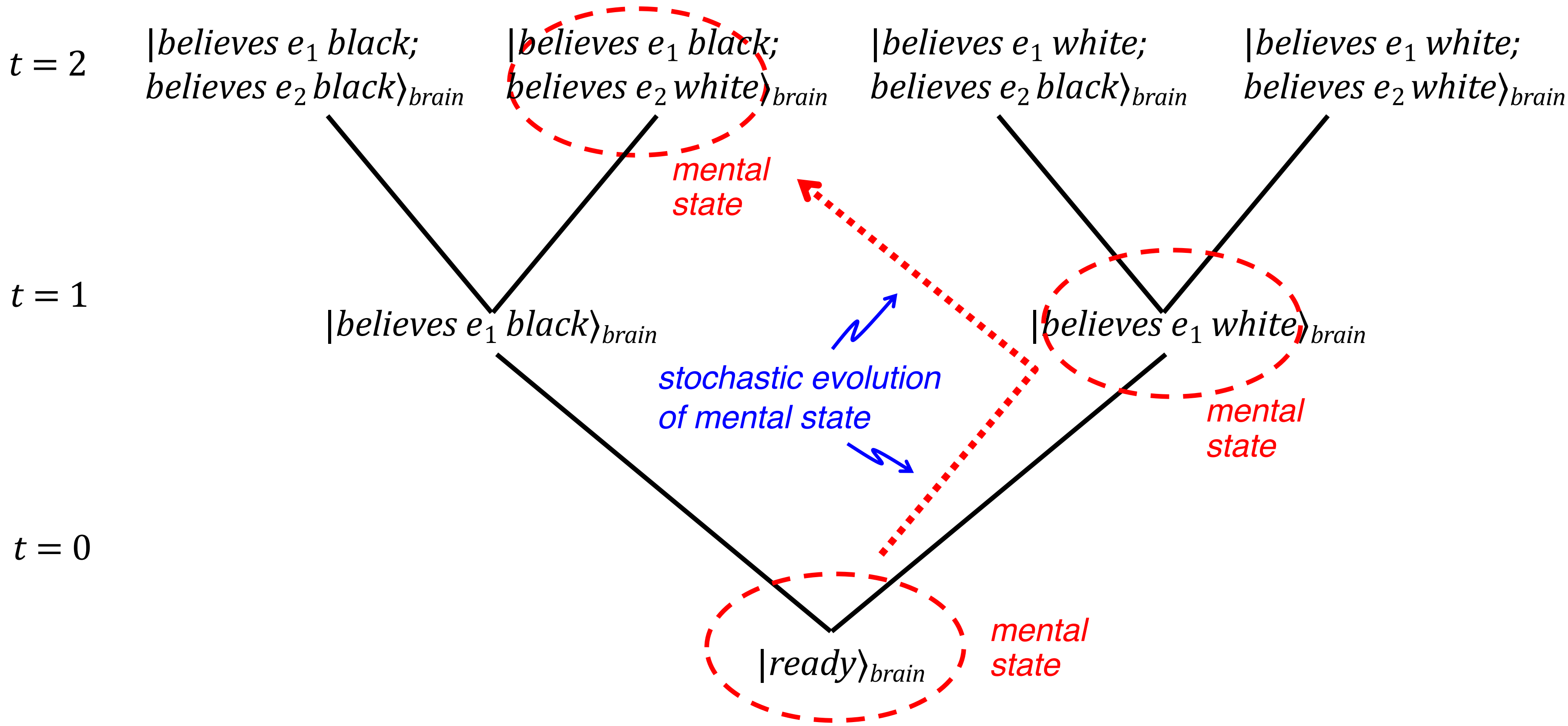
- *Motivation*: Suppose h conducts a *Color* measurement on e and ends up in standard state:

$$\sqrt{1/2}\,|\textit{believes e black}\rangle_h |\textit{"black"}\rangle_m |\textit{black}\rangle_e + \sqrt{1/2}\,|\textit{believes e white}\rangle_h |\textit{"white"}\rangle_m |\textit{white}\rangle_e$$

This is better than the Bare Theory:

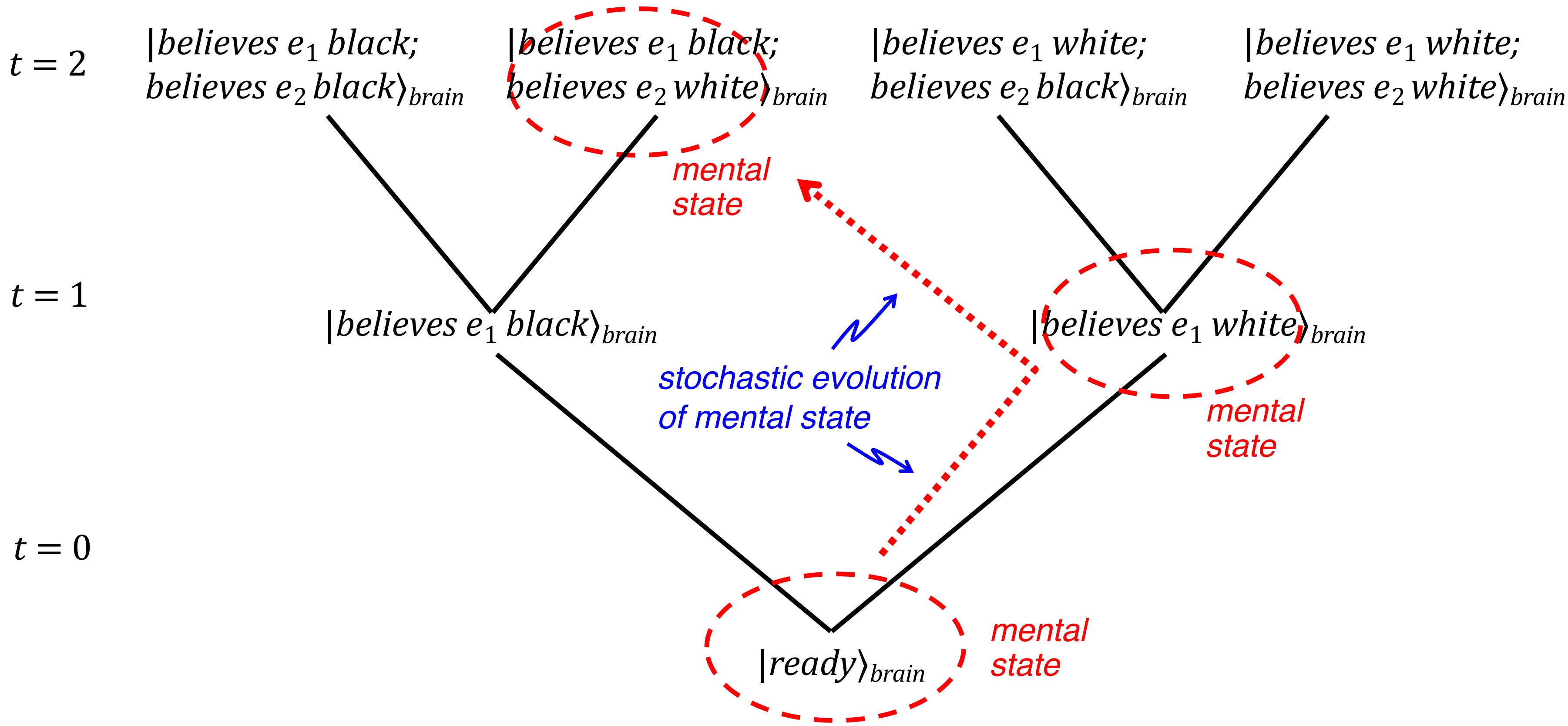
- *Bare Theory*: No explanation for our "effective knowledge" of measurement outcomes.
- *MM*: Our *mental states* explain our "effective knowledge" of measurement outcomes.
- *MM* entails we are not completely deceived by measurements: While our *brains* may be in superpositions, our *minds* are not!
- In particular, we will be correct about what we *report* we believe (even though our beliefs themselves will be incorrect). So common perceptions about introspection are upheld by *MM*.

How do mental states evolve?



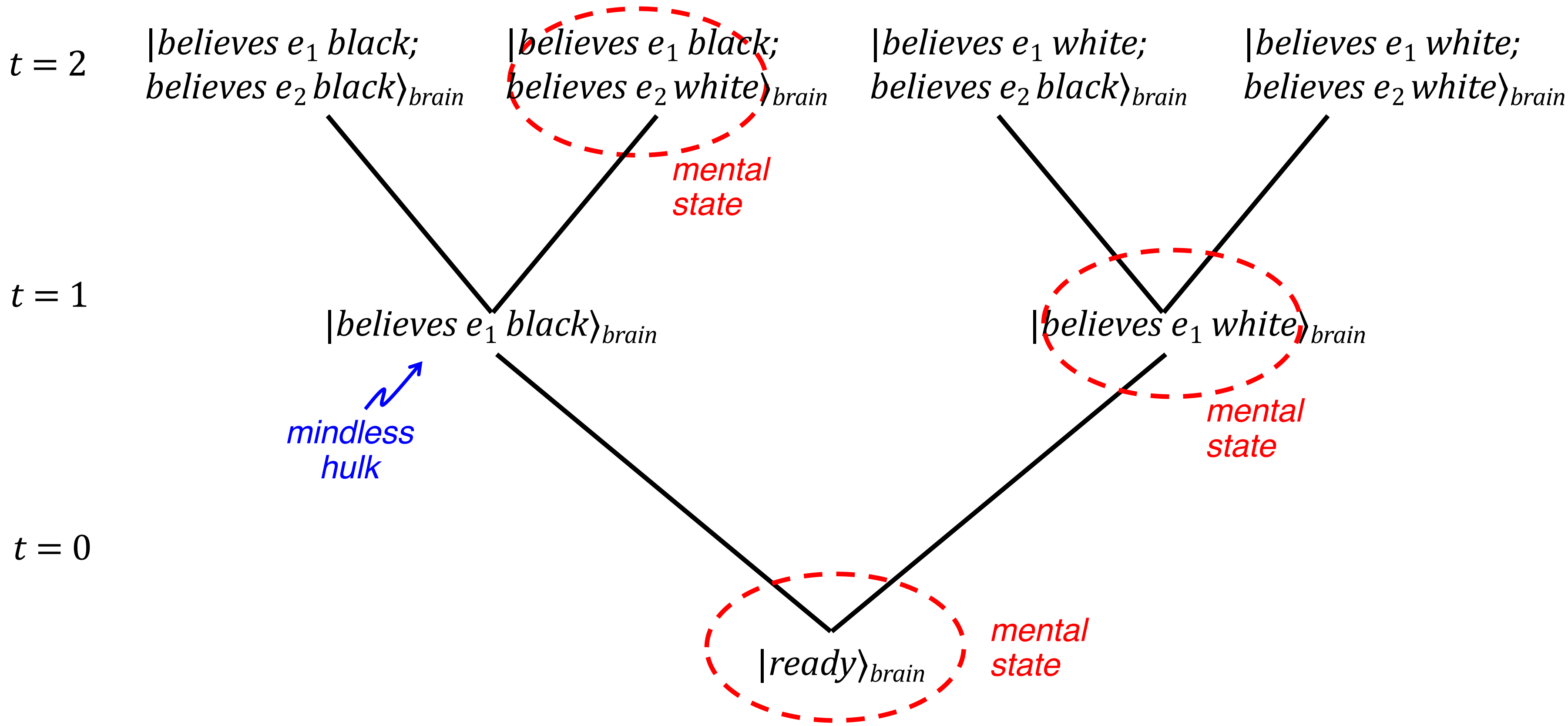
- Suppose h conducts a series of *Color* measurements on electrons e_1, e_2, ...
- *MM Claim*: h's mental state evolves stochastically in such a way that the probability of it being associated with a given brain state is given by the *Born Rule*.
- *After two measurements*:
  - h's brain state has evolved to a superposition with 4 terms.
  - h's mental state is associated with the belief that e_1 is black and e_2 is white.

How do mental states evolve?



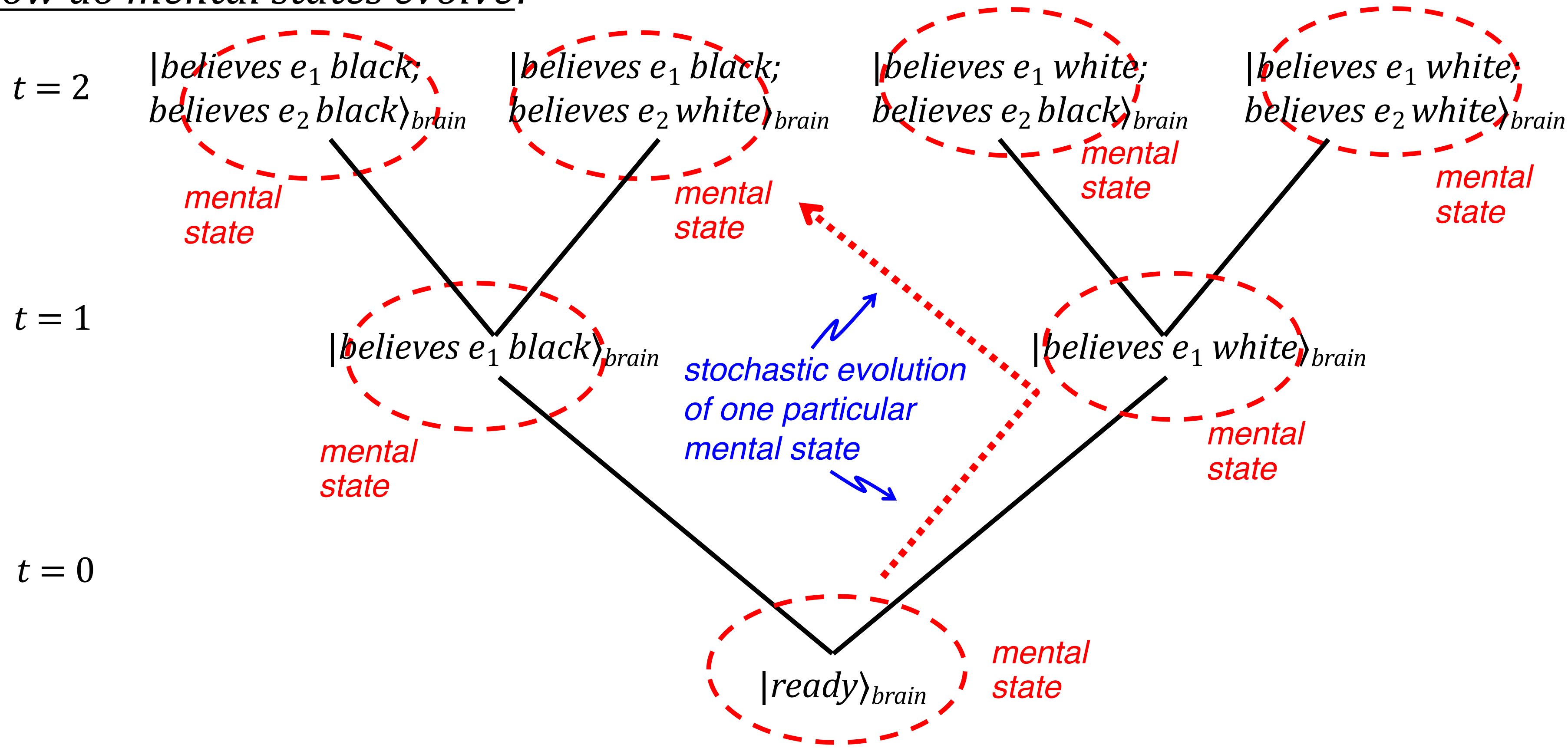
- <u>*So*</u>: h is mistaken about both her *current* and *past* beliefs about measurement outcomes:
  - At $t = 1$, she thinks she measured e_1 *white*; but at $t = 2$, she thinks she measured e_1 to be *black* at $t = 1$.
- <u>*But*</u>: At any given time, h's mental state will correspond to a definite brain state *that is not in a superposition*.
- <u>*So*</u>: h's beliefs *that* she has current and past beliefs are correct (*unlike* Bare Theory).

How do mental states evolve?



- *Initial Problem*: Only one of the terms in the superposition of brain states at any given time will be associated with a mental state.
- *Albert*: Most people we meet will be "mindless hulks"!

How do mental states evolve?



Remedy:

- Claim that h has a *continuous infinity* of minds (!!).
- Each individual mind evolves stochastically as before.
- The complete collection of all of h's minds (h's *global mental state*) evolves deterministically according to the Schrödinger dynamics (it get's divided up horizontally among the branches at any given time according to the *Born Rule*).

Problems with MM

(1) What's a mind?

- If there's an explicit distinction between mental states and physical states, why go to all the trouble of MM?
  - *Why not just use this distinction as a means of implementing the Projection Postulate?*

- *Recall*: The problem with reconciling the Projection Postulate with the Schrödinger dynamics was, in one form, determining just when the Projection Postulate applies.
  - *If we had a distinction between minds and bodies, we could simply say: Apply the Projection Postulate whenever a mind interacts with a body.*

Problems with MM

(2) How do probabilities appear in MM?

- MM, arguably, avoids the MW problems of preferred bases and conservation laws (how?), but what about the problem with probabilities?
- What is needed is an indeterministic dynamics of minds (as opposed to worlds) that agrees with the probabilities that quantum mechanics prescribes.
- *Albert*: These probabilities can simply be put in by fiat, stipulating that they agree with QM prescriptions.
- *But*: Is this an adequate response? Why can't MW respond in a similar way?