Dynamics By Itself, Everett, and Many Worlds

Christian Wüthrich

http://www.wuthrich.net/

MA Seminar: Philosophy of Physics

Hugh Everett III (1930-1982)

- entered graduate school at Princeton in 1953
- One night in 1954, "after a slosh or two of sherry," he had a conversation with fellow grad student Charles Misner and Aage Petersen (long-time assistant to Bohr) during which he had the basic idea behind the many-worlds theory
- began to work these ideas into a dissertation under the supervision of J.A. Wheeler
- Spring 1956: Wheeler takes draft to Copenhagen to discuss it with the Master, Pedersen, and Alexander Stern

Princeton, 1955: Everett (second from right) with Niels Bohr and Charles Misner (first from left).

The drama of the dissertation

- After these deliberations, Wheeler wrote back to Everett: "Your beautiful wave function formalism of course remains unshaken; but all of us feel that the real issue is the words that are to be attached to the quantities of the formalism."
- Here's a taste of the "words": "From the viewpoint of the theory, all elements of a superposition (all 'branches') are 'actual,' none any more 'real' than the rest." (In a footnote of the dissertation draft)
- In a letter to Stern, Wheeler excused Everett's theory as an extension, not a refutation of the received Copenhagen wisdom...

Wheeler to Stern

"I think I may say that this very fine and able and independently thinking young man had gradually come to accept the present approach to the measurement problem as correct and self-consistent, despite a few traces that remain in the present thesis of a past dubious attitude. So, to avoid any possible misunderstanding, let me say that Everett's thesis is not meant to question the present approach, but to accept it and generalize it."

From many worlds to Mutually Assured Destruction

- Of course, Everett would have completely disagreed!
- In 1957, Wheeler made Everett delete all unorthodox passages from his draft, cut it to one quarter of the original length, and submit the tamed version.
- In April 1957, the committee accepted the abridged version.
- Discouraged, he left academia to work on military and industrial mathematics and computation.
- In 1959-60, he helped to draft classified report WSEG No. 50 which overthrew prevailing nuclear military strategy by establishing the result of a nuclear conflict with the Soviet Union as Mutually Assured Destruction.

The original relative-state formulation

- Everett: what if the dynamical evolution of a quantum system is always in accordance with the Schrödinger eq?

$\Rightarrow$ no collapse, solves the MP by rejection of claim that there must be definite measurement results

$\Rightarrow$ doesn't rely on problematic distinction between micro- and macroworld or between object and observer or between conscious and non-conscious objects

- Instead: universal wave function, observer is inside the total system

$\Rightarrow$ Puzzle for Everett: how can we explain that the total system's (particle-apparatus-observer) being in a post-measurement state of entangled superposition of mutually incompatible records is in agreement with the empirical predictions made by standard QM?

- To this end, Everett presented a principle...

Principle (Fundamental Relativity of Quantum Mechanical States)

In the post-measurement superposition state, the observer records 'x-spin up' relative to the particle being in a state of x-spin up and 'x-spin down' relative to the particle being in a state of x-spin down.

- But this principle does not by itself provide the determinate measurement records predicted by standard QM.

$\Rightarrow$ gap between what Everett sets out to explain and what he delivers: "He set out to explain why observers get precisely the same sort of measurement records in his no-collapse formulation of quantum mechanics as predicted by the standard collapse formulation of quantum mechanics, but ends up describing a post-measurement observer who apparently does not have any particular measurement record." (Barrett 2008, SEP article on Everett, Sec. 3)

Three problems for the original relative-state theory

Barrett, Jeff (2008). Everett's relative-state formulation of quantum mechanics. Stanford Encyclopedia of Philosophy, http://plato.stanford.edu/entries/qm-everett/.

According to Barrett (2008), Everett's original relative-state theory suffers from three basic problems:

1. It offers no explanation of the sense in which the observer has, or appears to have, a determinate measurement record.
2. It fails to account for the standard probabilistic predictions of standard QM.
3. It is not empirically coherent, i.e. it doesn't explain how empirical justification for accepting it can be had when the world would in fact be faithfully described by it.

Various developments of Everett's original theory try to answer these three challenges, albeit in different contexts.

The dynamics by itself

Albert, Ch. 6.

The trouble with a purely linear dynamics was that the total post-color-measurement state of the observer o and measuring device m and the measured (previously hard) e^-

$$\frac{1}{\sqrt{2}}\left(|\text{`black'}\rangle_o|\text{`black'}\rangle_m|\text{black}\rangle_e + |\text{`white'}\rangle_o|\text{`white'}\rangle_m|\text{white}\rangle_e\right), \qquad (1)$$

i.e. there is no fact about what color the e^- has, or about what color the measuring device indicates, or about what the observer takes its color to be.

Everett announced that the post-color-measurement state of an initially hard e^- is, in fact, just the superposition like the one in (1)...

Problems: probabilities, classicality/basis-dependence

1. How do we get the right probabilities?
2. Albert: "The trouble is that what worlds there are... will depend on what separate terms there are in the universal state vector at that instant; and what separate terms there are in that state vector at that particular instant will depend on what basis we choose to write that vector down in." (113)
    - But there is no preferred basis that would do this for us—at least not on the standard way of thinking...

- Suppose post-measurement state in (1) obtains.

⇒ It seems as if there's a world in which e$^-$ is black and one in which it is white.

- But note that (1) can be expressed in a different basis:

$$\frac{1}{\sqrt{2}}\left(|Q^+\rangle_{(h\&m)}|\text{hard}\rangle_e + |Q^-\rangle_{(h\&m)}|\text{soft}\rangle_e\right),$$

where

$$\begin{aligned}
|Q^+\rangle_{(h\&m)} &= \frac{1}{\sqrt{2}}\left(|\text{`black'}\rangle_o|\text{`black'}\rangle_m + |\text{`white'}\rangle_o|\text{`white'}\rangle_m\right), \\
|Q^-\rangle_{(h\&m)} &= \frac{1}{\sqrt{2}}\left(|\text{`black'}\rangle_o|\text{`black'}\rangle_m - |\text{`white'}\rangle_o|\text{`white'}\rangle_m\right).
\end{aligned}$$

⇒ But in this basis it seems as if there's a world in which e$^-$ is hard and one which it is white...

Dynamics by itself: the 'bare theory'

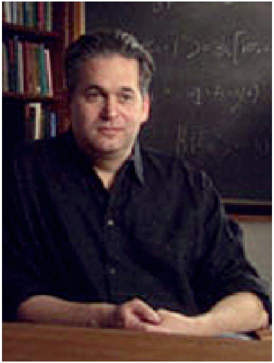
- Question: what would it be like to be in a state such as (1)?
- Let's consider what an observer in state (1) would respond to question.
- No good to ask "what do you presently believe is the color of the e^-", since the state of the world would be in superposition after her response. (Why?)

Fact

The linearity of those operators that represent observables ('linear operators') implies that an observable $\mathcal{O}$ of any system S has the same determinate value in the state $|A\rangle_S$ and in the state $|B\rangle_S$, then $\mathcal{O}$ also has that same determinate value in any linear superposition of $|A\rangle_S$ and $|B\rangle_S$. (cf. 117)

- Apply this Fact to the state in (1) and prompt the observer to tell us merely whether or not the e^- has one of the two colors black and white.
- She will indeed affirm this if she is a competent reporter: if $|\text{'black'}\rangle_o|\text{'black'}\rangle_m|\text{black}\rangle_e$ obtains, then she has a definite belief and will report that one of those is indeed the case; analogously for $|\text{'white'}\rangle_o|\text{'white'}\rangle_m|\text{white}\rangle_e$.
- By the Fact, her brain will thus be in a definite state if the superposition in (1) obtains and she will report so.
- Odd: the linear dynamics seems to entail that she is going to be convinced that she has a definite particular belief about the color of the e^-!

$\Rightarrow$ If a state like (1) obtains, then she seems to be "deceived even about what her own occurrent mental state is." (118)



"And so it turns out that there was... too much being taken for granted when we got convinced... that there is some particular point in the course of the sort of measurement we've been talking about by which a collapse of the wave function must necessarily already have taken place, some particular point... at which the dynamical equations of motion together with the standard way of thinking about what it means to be in a superposition somehow flatly contradicts what we unmistakably know to be true of our own mental lives." (118f)

Suppose the observer carries out two subsequent color measurements ($m1$ and $m2$) on an initially hard e^-. Assuming her competence, the post-measurement state of the composite system ($o\&m1\&m2\&e$) is

$$1/\sqrt{2}\ (|\text{both outcomes are `black'}\rangle_o |\text{`black'}\rangle_{m1} |\text{`black'}\rangle_{m2} |\text{black}\rangle_e + |\text{both outcomes are `white'}\rangle_o |\text{`white'}\rangle_{m1} |\text{`white'}\rangle_{m2} |\text{white}\rangle_e) \quad (2)$$

The questions to ask the observer then are (1) whether she believes that both measurements had definite outcomes, and if so, (2) whether they were the same.

Given the previous argument, she will affirm both these questions.

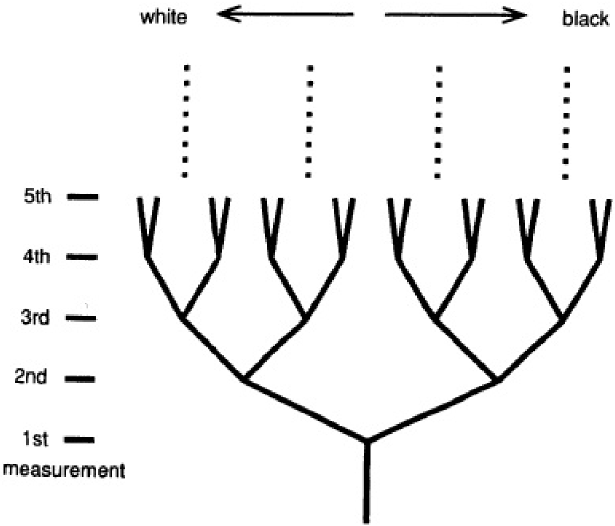
"Effective knowledge"

- From the sorts of arguments, we can see that if two different observers were to perform subsequent color measurements on the same particular initially hard e^-, "then both of those observers will report, falsely, that the other observer has reported some definite particular outcome of her measurement, and both of them will report that that reported outcome is completely in agreement with her own." (120)

$\Rightarrow$ If a state like (1) obtains, then the observer "effectively knows" what the color of the e^- is.

- Let's do the same spiel for an observer who performs color measurements on an infinite collection of initially hard e^-... (but let's do this on the blackboard)

An infinite sequence of color measurements



Getting the statistics back

- Ask the infinite experimenter another question: "Tell me whether you believe that each of the first N e^- has a definite color, and if so, tell me what fraction of them do you believe to be black."
- The observer "effectively knows" the color of the first N e^- and will thus answer 'yes' to the first Q.
- But the second Q is more complicated, since the world will be in a superposition of states with the observer giving different answers to it.
- However, in the long run, as N goes to infinity, "the state of the world will, with certainty, approach a state in which [the observer] will answer that question in a perfectly determinate way, and in which the answer [she] gives will with certainty be '1/2'." (122)
- It can be shown that the pre-measurement state of the collection of e^- is an eigenstate of the operator corresponding to the observable $\mathcal{O}_N$ (the fraction of black e^- among the first N measured e^-) as N goes to infinity, with eigenvalue 1/2. (Cf. fn. 6, p. 122)
- But since it's an eigenstate of the operator corresponding to the measurement, the post-measurement state will determinately have 1/2 black e^-.

The upshot of all of this

> *"Suppose that an observer h is confronted with an infinite ensemble of identical systems in identical states and that she carries out a certain identical measurement on each one of them. Then, even though there will actually be no matter of fact about what h takes the outcome of any of those measurements to be, nonetheless as those measurements which have already been carried out goes to infinity, the state of the world will approach (as a well-defined mathematical limit) a state in which the report of h about the statistical frequency of any particular outcome of those measurements will be perfectly definite, and also perfectly in accord with the standard quantum-mechanical predictions about what that frequency ought to be." (123)*

In other words, we can get the correct statistics without resorting to collapses!

A problem for the bare theory: infinity

- Problem: these frequencies will only be factual in the limit of infinitely many measurements. There simply are no matters of fact about these frequencies otherwise...
- But of course we could only perform an infinite series of measurements if we had an infinite amount of time, or if there were infinitely many of us.

⇒ There is no matter of fact "about whether or not we take those frequencies to be in accordance with the standard quantum-mechanical predictions about them." (124)

⇒ What sorts of reason could we then have to believe in QM in the first place?

⇒ Albert concludes that the bare theory is not "entertainable", i.e. it can't be the case that the linear Schrödinger evolution gives the true and complete dynamics.

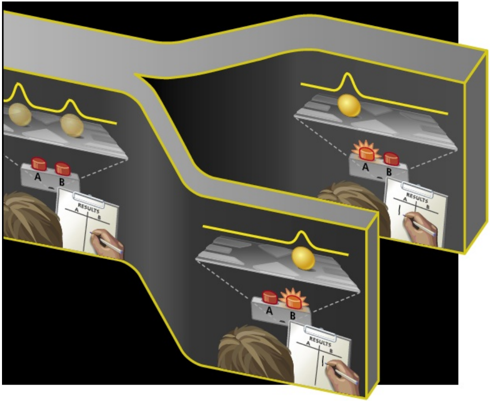
Developments of the bare theory

So given that the bare theory is, in Barrett's words, not empirically coherent, just like Everett's original formulation, let's look for ways to develop Everett's rejection of the collapse postulate and see where that leads us. Options:

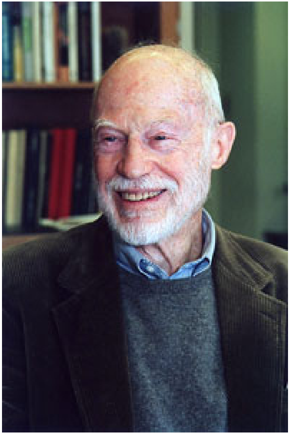
1. canonical many-worlds interpretation of DeWitt
2. many minds (Albert, Loewer)
3. many histories

(1) Many worlds (Bryce DeWitt)

DeWitt 1970, 1971, 1973.



- wave function describing system bifurcates at each interaction of observer with superposed object
- no interaction between branches (which each contains complete copy of system)
- today's explanation of how branches become independent: decoherence theory



Bryce DeWitt (1923-2004)

Of course, DeWitt's splitting of worlds whenever the states of systems become correlated is counterintuitive, as he freely admits:

> *"I still recall vividly the shock I experienced on first encountering this multiworld concept. The idea of 10^{100} slightly imperfect copies of oneself all constantly splitting into further copies, which ultimately become unrecognizable, is not easy to reconcile with common sense. Here is schizophrenia with a vengeance." (1973, 161)*

Problems of the many-worlds view

1. interpretational: very likely not how Everett conceived of his proposal, since Everett's original suggestion did not involve physical splitting of observers or other physical systems
2. ontological extravaganza
3. preferred-basis problem: this was already addressed as one of Albert's points of critique; it is a serious problem, and it's unclear as of yet whether there are completely satisfactory answers to it; most people in this camp use decoherence theory to solve this problem, i.e. tell a story about how the interaction of the system with the environment (settings on apparatus, etc) determines the basis along which splitting occurs.

4. getting the statistics right: also already discussed; there exist various proposals to resolve this issue, and it's debated whether they solve the problem; what is clear, though, is that Everett's own proposal of conceiving of these probabilities in the same way as in classical thermodynamics doesn't work without further assumptions.
5. potential incompatibility with special relativity: particularly those who maintain that the splittings are physical have great trouble reconciling it with SR; this is a deep and completely unresolved issue.

(2) Many minds (Albert and Loewer)

- Given the problems of the bare theory and the canonical many-worlds interpretations, Albert and Loewer (1988) propose an alternative way of solving the MP by denying Maudlin's 1.C: the many-minds interpretation.
- Suppose that there's just one world with one complete and true story of how it is.
- Assume completeness (1.A) and linear dynamics (1.B).
- Suppose that healthy people are competent reporters of whether they have determinate mental states (such as 'the position of the pointer is such and such')

⇒ Dynamical equations entail that healthy people in superpositions of brain states corresponding to different beliefs about pointer positions will report (with certainty) that they are in determinate mental states about the position of the pointer...

⇒ "something funny about how mental states supervene on brain states" (126)

Let's have a closer look

- evolution of brain states in accordance with deterministic Schrödinger eq, but evolution of mental states is probabilistic
- Observer starts out in $|\text{ready}\rangle_o$, her brain state will be in superposition of $|\text{'black'}\rangle_o$ and $|\text{'white'}\rangle_o$, but her post-measurement mental state will be either in the mental state corresponding to $|\text{'black'}\rangle_o$ or in the mental state corresponding to $|\text{'white'}\rangle_o$ (with equal probabilities)

"Whatever belief [she] does end up with, when [(1)] obtains, is necessarily going to be a false belief. But there are very natural ways of cooking things up so as to guarantee that that belief will nonetheless have an important kind of effective validity, at least in so far as [she] is concerned... there are ways of cooking things up... so as to guarantee that the future evolution of [her] mental state will proceed... exactly as if [her] beliefs were true." (127)

Here's what this means:

- Suppose she performs two subsequent color measurements on hard e^-.
- After the first measurement, she has the determinate mental state of believing that the first outcome was 'black' (or 'white').
- After the second measurement, the physical state will be (2), and she will end up with determinate belief that both outcomes were 'black' (or both outcomes were 'white').
- But that's precisely how her mental state would have been, with certainty, had her belief that the first outcome was 'black' been true.
- And this is true quite generally.

"On this proposal, quantum-mechanical wave functions are complete descriptions of the physical state of things, and those wave functions invariably evolve in perfect accordance with the dynamical equations of motion, and it makes no physical difference at all what basis we choose to write those wave functions down in, and measurements carried out by sentient observers (that is: by observers with minds) invariably have determinate outcomes in the minds of those observers, and the statistical distributions of those outcomes will be the usual quantum-mechanical ones, and there isn't anything mysterious about how probabilities come up in the theory, and the reports of sentient observers about their own mental states will invariably... be correct." (129, footnotes suppressed)

Some consequences...

- severe form of dualism: all but one of the terms in a superposition like (1) represents "mindless hulks" (and which of them isn't mindless can't be inferred from complete knowledge of physical state)
- Fix: "every sentient physical system there is is associated not with a single mind but rather with a continuous infinity of minds..." (130)
- dynamical evolution of minds is probabilistic such as to yield the correct QM-statistics
- evolution of the continuously many minds connected to one sentient observer as a set always evolves deterministically

$\Rightarrow$ On this proposal, the 'global' states of sentient observers is uniquely determined by the physical state of the world.

- There's no physically preferred basis, only a mentally preferred basis.
- It is one of the few interpretations of QM that is manifestly compatible with SR (but only in physical realm).
- It can be thought of as an AV theory, with the mental states as additional variables.

⇒ This is sufficient to give observers determinate records.

- It is a bit as if all problems (preferred basis, compatibility with SR, creation of determinate records, etc.) are simply pushed into the mental realm...

One more thing:

Physics, on this proposal, is local in some sense, in apparent contradiction to Bell's theorem. You should have a close look at this, considering a EPR-type state.

Weaknesses and problems

This proposal has some weaknesses and problems:

- It entails that the beliefs of sentient observers will typically be false. Albert: "Nothing, even in principle, can be done about that." (132)
- The "sum total of what any particular... observer's minds can conclude about the overall quantum state of the world", given the outcomes of any experiments she might perform, is only that the state of the world "is not orthogonal... to the effective state that those outcomes pick out." (133)
- But the strongest problem, as far as most commentators are concerned (such as Barrett 2008), is the strong mind-body dualism that is implied by the view. Barrett wonders "whether the sort of mental supervenience one gets is worth the trouble of postulating a continuous infinity of [non-physical] minds associated with each observer." (2008, Sec. 6)