The Many-Worlds and Many-Minds Formulations of Quantum Mechanics

Jeffrey A. Barrett
4 January 2004

The many-worlds and many-minds formulations of quantum mechanics are reconstructions of Hugh Everett III's (1957a, 1957b, 1973) relative-state formulation of quantum mechanics. Each is presented as a proposal for solving the quantum measurement problem. Much of the philosophical interest in these theories derives from the metaphysical commitments they suggest. They illustrate the roles played by traditional metaphysical distinctions both in formulating and in evaluating our best physical theories. They also illustrate the range of metaphysical options one must consider if one wants a metaphysics that is consistent with the structure of the physical world suggested by our best physical theories.

The quantum measurement problem is a consequence of the orthodox quantum-mechanical representation of physical properties. In order to account for interference effects, the orthodox view requires that we allow for a physical system to be in a *superposition* of having mutually incompatible classical physical properties. An electron *e* might, for example, be in a superposition of being in New York City and being in Los Angeles. If the unit-length vector |NYC>_e represents the electron being in New York City and if the orthogonal unit-length vector |LA>_e represents the electron being in Los Angeles, then the state of the electron in a superposition of being in each city is represented by

|S>_e = a|NYC>_e + b|LA>_e,

where *a* and *b* are complex numbers, such that *a*-squared plus *b*-squared equals one. On the orthodox view, the state represented by the unit-length vector |S|>_e is not a state where the electron is determinately in NYC, it is not a state where the electron is determinately in LA, it is not a state where the electron is determinately in both cities, and it is not a state where the electron is determinately in neither city. Rather, on the standard interpretation of states, an electron in state |S>_e simply fails to have a determinate position.

While allowing for superpositions of classical properties explains the counterintuitive empirical results of interference experiments, it leaves us with a puzzle: If electrons are sometimes in such superpositions of position, then why do we always find electrons to have determinate positions whenever we look for them? In its most general form the quantum measurement problem is to explain why physical systems exhibit quantum interference effects, which typically involves talk of superpositions, *and* to explain why we always observe them to have determinate physical properties when we look.

The standard von Neumann-Dirac collapse formulation of quantum mechanics (1932) explains interference effects and definite measurement results by stipulating two dynamical laws. The *linear dynamics* describes the deterministic continuous evolution of the state of a physical system when no measurement of the system is made. It is this law that describes the evolution of physical systems in superpositions of classical properties and thus explains quantum interference effects. The *collapse dynamics* describes the random discontinuous evolution of the state when a measurement is made of the physical system. It is this law that explains how we get determinate measurement records at the end of an observation and makes the standard statistical predictions. More specifically, in the case of the electron in state |S>_e, if an observer *M* looks for the electron in NYC, the collapse dynamics predicts that the state will instantaneously and randomly evolve from

|Ready>_M (a|NYC>_e + b|LA>_e),

a state where M is ready to look for *e* and *e* is in state |S>_e, to either |"In NYC">_M |NYC>_e (with probability |a|^2), in which case *e* is now determinately in NYC and *M* determinately records this fact, or to |"Not in NYC">_M |LA>_e (with probability |b|^2), in which case *e* is now determinately in LA and *M* determinately records that it is not found in NYC.

In order to understand the work done by the collapse dynamics in the standard theory, consider what would happen without the collapse of the quantum-mechanical state. In the measurement above, the linear dynamics predicts that the post-measurement state of an observer who correlates his records perfectly with the position of the electron, written in the determinate record basis, is

|E> = a|"In NYC">_M |NYC>_e + b|"Not in NYC">_M |LA>_e.

On the standard interpretation of states, *M* here has no determinate measurement record. Rather, without the collapse dynamics, *M* ends up in an entangled superposition of finding and not finding the electron. This is presumably not what happens.

So in the standard theory, the collapse dynamics is both responsible for the theory making the standard quantum statistical predictions and for the explanation of determinate measurement results. But since the physical state that results from applying the collapse dynamics to a system is typically different from the state that results from applying only the linear dynamics, the standard formulation of quantum mechanics is at best incomplete and arguably logically inconsistent on a strict reading unless one can stipulate strictly disjoint conditions for the when each dynamical laws obtains. In the context of the standard collapse theory solving the measurement problem would require on to stipulate exactly what interactions count as measurements and hence cause collapses.

Rather than stipulating when collapses occur, Everett's proposal for solving the quantum measurement problem involved denying that are collapses. More specifically Everett proposed simply dropping the collapse dynamics from the standard von Neumann-Dirac theory of quantum mechanics and taking the resulting pure wave mechanics as a complete and accurate description of all physical systems. Everett then intended to deduce the standard statistical predictions of quantum mechanics, the predictions that are explained by the collapse dynamics in the standard formulation of quantum mechanics, as subjective experiences of observers who are themselves treated as ordinary physical systems within the new theory. Dropping the collapse dynamics clearly eliminates potential conflict between the two dynamical laws; but if one drops the collapse dynamics, one must then explain how we get determinate measurement results that exhibit the standard quantum statistics since the linear dynamics alone typically predicts entangled post-measurement superpositions like |E>_e.

While it is clear that Everett intended for his relative-state formulation of quantum mechanics to explain why we get determinate measurement results, it is unclear how this was supposed to work. There are several alternative reconstructions of Everett's theory in the literature, all designed to provide quantum mechanics without the collapse dynamics with determinate measurement records while somehow recovering the standard quantum statistics. The many-worlds and the many-minds formulations of quantum mechanics represent two general approaches to reconstructing Everett's relative-state formulation of quantum mechanics.

The splitting-worlds theory is perhaps the most popular version of the many-worlds formulation. The splitting-world formulation of quantum mechanics

> asserts that it makes sense to talk about a state vector for the whole universe. This state vector never collapses and hence reality as a whole is rigorously deterministic. This reality, which is described *jointly* by the dynamical variables and the state vector, is not the reality we customarily think of, but is a reality composed of many worlds. By virtue of the temporal development of the dynamical variables the state vector decomposes naturally into orthogonal vectors, reflecting a continual splitting of the universe into a multitude of mutually unobservable but equally real worlds, in each of which every good measurement has yielded a definite result and in most of which the familiar statistical quantum laws hold." (DeWitt and Graham 1973, v)

Proponents of this view admit that the metaphysical commitments it suggest are counterintuitive: "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 spitting into further copies, which ultimately become unrecognizable, is not easy to reconcile with common sense. Here is schizophrenia with a vengeance." (1973, 161)

But it is precisely these counterintuitive commitments that explain why observers end up recording determinate measurement results. On the splitting-worlds formulation the universe splits whenever one makes a measurement in such a way that every physical possible result in fact determinately occurs in some future world. More specially, there is one world corresponding to each term in the expression of the quantum mechanical state when written in the theory's preferred basis. In choosing the preferred basis, one chooses a single preferred way from among the many different, mathematically equivalent, ways of representing quantum-mechanical states as the sum of mutually orthogonal unit-length vectors. On the splitting-worlds formulation, the preferred basis is chosen so that each term in the expansion of the state describes a world where there is a determinate measurement record. The state |E> above describes two worlds: One where the observer *M* determinately records the measurement result "In NYC" and *e* is in fact in NYC and another where *M* determinately records "Not in NYC" and *e* is in fact in LA.

While the splitting-worlds formulation of quantum mechanics does explain why we have determinate measurement records, it encounters other problems. A standard complaint is that the theory is ontologically extravagant. We presumably only ever need one physical world, *our* world, to explain for our experiences. The reason for postulating the actual existence of a different physical world corresponding to each term in the quantum-mechanical state is that is allows one to explain our determinate experiences while taking the deterministically-evolving quantum-mechanical state to be in some sense a complete and accurate description of the physical facts. But again one might wonder whether the sort of completeness one gets warrants the many-world ontology.

Another problem with the splitting-worlds formulation concerns statistical predictions of future events. The standard collapse formulation of quantum mechanics predicts that *M* will get the result "In NYC" with probability *a*-squared and the result "Not in NYC" with probability *b*-squared in the above experiment, and this is what we observe as relative frequencies for such experiments. Insofar as their will be two copies of *M* in the future, *M* is guaranteed to get each of the two possible measurement results. So, in this sense at least, the probability of *M* getting the result "In NYC" is one, which is simply not what we observe. But that is the wrong answer. A principle of indifference might lead one to assign probability ½ to each of the two possible measurement outcomes. But not only would such a principle be difficult to justify here, probability ½ for each possible outcome is typically not what we observe for such experiments as relative frequencies. So while the splitting-worlds formulation explains why observers get determinate measurement records, it makes no empirical predictions for the likelihood of future events.

In order to understand what one would have to add to the theory to get the standard quantum statistical predictions for future events, one might note that the question "What is the probability that *M* will record the result 'In NYC'?" is, strictly speaking, nonsense unless one has an account of the transtemporal identity of the observer *M*. It is the fact that there is no rule telling us which worlds are which at different times that prevents the splitting-worlds theory from making statistical predictions concerning an observer's future experiences. And not being able to account for the standard quantum probabilities is a serious problem since it was the successful statistical predictions of quantum mechanics that made quantum mechanics worth taking seriously in the first place.

Another problem for the splitting-worlds formulation of quantum mechanics concerns the way worlds are supposed to split. In order to explain our determinate measurement records, one must choose a preferred basis so that observers have determinate measurement records in each term of the quantum-mechanical state when written in the preferred basis. The problem is that not just any basis will make records determinate in every world (consider, for example, a basis that includes the vector |E> above). Selecting the preferred basis to use determines when worlds split, and determining when worlds split is as difficult as trying to determine when the collapse occurs in the standard formulation of quantum mechanics. This is the preferred basis problem. This problem is closely analogous to the original measurement problem in the context standard collapse formulation of quantum mechanics.

A popular strategy for resolving the preferred basis problem is to try to find a criterion involving the interaction between a quantum-mechanical system and its environment that would dynamically select a preferred basis for a system. As a simple example of an environmental decoherence criterion, one might take the preferred basis of a system to be the one that represents the classical property of the system to which its environment becomes most strongly correlated. Insofar as measurement records are easily read, their environments become strongly corralled with them, so such a criterion would can be expected select the determinate-record basis as preferred. One problem with having the environment of a system select the preferred basis, however, is that, at least here, one presumably needs a preferred basis for the entire universe, which does have an environment.

David Albert and Barry Loewer's many-minds formulation of quantum mechanics (1988) provides another approach for interpreting Everett's relative-state formulation of quantum mechanics. Everett said that his theory "is objectively continuous and causal, while subjectively discontinuous and probabilistic" (1973, 9). The many-minds formulation of quantum mechanics captures this feature by distinguishing between an observer's *physical state* and its evolution, which is continuous and causal, and an observer's *mental state* and its evolution, which is discontinuous and probabilistic. This is a sort of hidden-variable theory, where the variable being added to the standard quantum-mechanical state is the mental states of observers. Stipulating determinate mental states solves the quantum measurement problem by directly providing observers with determinate, accessible measurement records.

In order to get the observer's mental state to supervene on his physical state, Albert and Loewer associate with each observer a continuous infinity of minds. The standard quantum-mechanical state always evolves in the usual deterministic linear way, but each mind evolves randomly, with probabilities determined by the particular mind's current mental state and the evolution of the quantum-mechanical state. In the experiment above, Albert and Loewer's mental dynamics predicts that the probability of each of the observer's minds becoming randomly associated with the result "In NYC" (the first term of |E>) is *a*-squared and that the probability of each becoming randomly associated with the result "Not in NYC" (the second term of |E>) is *b*-squared.

An advantage of the many-minds formulation over the splitting-worlds formulation is that here there is no physically preferred basis. One must choose a preferred basis in order to specify the mental dynamics completely, but this choice has nothing to do with any physical facts. Rather, it can be thought of as part of the description of the relationship between physical and mental states. Another advantage of the many-minds formulation is that, unlike the splitting-worlds formulation, it makes the standard probabilistic predictions for the future measurement results of each mind. Since the states of particular minds do not supervene on the physical state here, in order to talk about their states and how they evolve, one must suppose that individual minds have transtemporal identities, which in turn requires a commitment to a strong form of mind-body dualism. But it is also this strong dualism that makes the many-minds theory one of the few formulations of quantum mechanics that resolves the quantum measurement problem and is manifestly compatible with special relativity.

One might wonder whether the sort of mental supervenience one gets in the many-mind formulation (it is not the states of an observer's individual minds, but only the *distribution* of the states of all of these minds that can be taken to supervene on her physical state) is worth the trouble of postulating a continuous infinity of minds associated with each observer. Another option is to suppose that each observer has a *single* mind that evolves in the Albert and Loewer random way. But here one gives sacrifices all but the weakest sort of supervenience of mental states on physical states. Here the physical state would only tell one the probabilities of various mental states obtaining.

If one wants to avoid the mind-body dualism involved in the many-minds formulation, one can use the evolution of minds to construct an alternative many-worlds theory. On one such theory, the many-threads formulation of quantum mechanics, worlds do not split. Rather, one stipulates that there is one world corresponding to each possible trajectory of a single Albert-Lower mind and that the history of that world is described by the history of the world that would be observed by the mind and that we in fact inhabit exactly one of these worlds. The global quantum-mechanical can be used to assign a prior probability to each physically possible history describing the history of *our* world. These prior probabilities, concerning which possible world is our world, might then be updated as one learns more about the history of our world. In the simplest case, when an event occurs in our world, one can eliminate from contention all worlds that are incompatible with the event. Unlike the splitting-worlds formulation, there is no special problem here in understanding probabilities of future events. A particular event is either going to happen or not in our world. The standard quantum probabilities here simply represent our posterior uncertainly concerning which world we in fact inhabit (Barrett 1999).

Bibliography and Suggested Reading

Albert, D. Z: 1992, *Quantum Mechanics and Experience*, Harvard University Press, Cambridge, MA.

Albert, D. and B. Loewer: 1988, "Interpreting the Many Worlds Interpretation," *Synthese* 77: 195-213.

Barrett, J.: 1999, *The Quantum Mechanics of Minds and Worlds*, Oxford University Press, Oxford.

DeWitt, B. S.: 1971, "The Many-Universes Interpretation of Quantum Mechanics," in *Foundations of Quantum Mechanics*, Academic Press, New York; reprinted in DeWitt and Graham (eds), 167-218.

DeWitt, B. S. and N. Graham (eds): 1973, *The Many-Worlds Interpretation of Quantum Mechanics*, Princeton University Press, Princeton.

Dowker, F. A. and Kent: 1996, "On the Consistent Histories Approach to Quantum Mechanics," *Journal of Statistical Physics*, vol. 83, nos. 5-6, 1575-1646.

Everett, H: 1957a, *On the Foundations of Quantum Mechanics*, thesis submitted to Princeton University, March, 1, 1957, in partial fulfillment of the requirements for the Ph.D. degree.

Everett, H: 1957b, "'Relative State' Formulation of Quantum Mechanics," *Reviews of Modern Physics*, 29: 454-462.

Everett, H: 1973, "The Theory of the Universal Wave Function," in DeWitt and Graham (eds).

von Neumann, J.: 1955, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, Princeton; translated by R. Beyer from *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin, 1932.