Brit. J. Phil. Sci. **56** (2005), 615–634

A Philosopher Looks at Quantum Mechanics (Again)*

Hilary Putnam

ABSTRACT

'A Philosopher Looks at Quantum Mechanics' (Putnam [1965]) explained why the interpretation of quantum mechanics is a philosophical problem in detail, but with only the necessary minimum of technicalities, in the hope of making the difficulties intelligible to as wide an audience as possible. When I wrote it, I had not seen Bell ([1964]), nor (of course) had I seen Ghirardi et al. ([1986]). And I did not discuss the 'Many Worlds' interpretation. For all these reasons, I have decided to make a similar attempt forty years later, taking account of additional interpretations and of our knowledge concerning non-locality. (The Quantum Logical interpretation proposed in Putnam [1968] is not considered in the present paper, however, because Putnam [1994b] concluded that it was unworkable.) Rather than advocate a particular interpretation, this paper classifies the possible *kinds* of interpretation, subject only to the constraints of a very broadly construed scientific realism. Section 7 does, however, argue that two sorts of interpretation—ones according to which a 'collapse' is brought about by the measurement (e.g. the traditional 'Copenhagen' interpretation), and the Many Worlds interpretation or interpretations—should be ruled out. The concluding section suggests some possible morals of a cosmological character.

1 Background

Forty years ago, I published a paper titled 'A Philosopher Looks at Quantum Mechanics' (Putnam [1965]) that many people found useful as an introductory

* This paper originated as a keynote address to the Annual Conference of the Israel Association for History and Philosophy of Science at the Van Leer Foundation in Jerusalem on 17 March 2004.

doi:10.1093/bjps/axi135

Advance Access published on September 9, 2005.

explanation of the controversies about the interpretation of quantum mechanics. I have been aware for many years of the need to return to the subject because so much relevant theorizing was not known when I wrote that paper. For example, 'A Philosopher Looks at Quantum Mechanics' was written in 1963–4, and I had not seen Bell's famous paper on the Einstein–Podolsky–Rosen paradox (Bell [1964]). Bell's central claim was that, if quantum mechanics is right, the measured values of spin on certain pairs of separated particles (electrons[1] in an 'entangled' state) would be incompatible with the classical postulate of 'locality'. Simply put, what 'locality' means is that these experiments could be set up in such a way that the measurement of the spin of particle 1 produces no physical disturbance in particle 2. As David Albert explains in an excellent introduction to the topic, locality, in this sense, seemed almost self-evident. There seemed to be any number of ways you could do it: you could, for example, separate the two particles by some immense distance so large that there is no time for a 'signal' from one of the particles to reach the other without travelling faster than light, or build impenetrable shields between them, or 'set up any array of detectors you like in order to verify that no measurable signals ever pass from one of the electrons to the other in the course of the experiment (since quantum mechanics predicts that no such array, in such circumstances, whatever sorts of signals it may be designed to detect, will ever register anything)' (Albert [1992], p. 64). The experiment described by Aspect et al. ([1982]), showed that that the 'non-locality' that Bell had derived from quantum mechanics in 1964 really exists, and ever since the question of how to understand non-locality has been at the very centre of discussions of the interpretation of quantum mechanics.[2]

In addition, at least one major interpretation of quantum mechanics that I will discuss in this essay had not yet been proposed. In brief, we now know many more facts and we also know of many more possibilities of interpretation than I did in 1965. Moreover, I failed to discuss in 'A Philosopher Looks at Quantum Mechanics' the 'Many Worlds' interpretation proposed

[1] However, non-locality can also be demonstrated with experiments involving other magnitudes than spin and other particles than electrons, and also with experiments involving fields rather than particles. See Finkelstein ([1987]) for a very clear explanation of non-locality using photons as the particles and direction of polarization as the relevant observable.

[2] Most laypeople, when they hear that quantum mechanics predicts non-local correlations, immediately conclude that this shows that 'causal signals can travel faster than light', thus contradicting Einstein's theories of Special and General Relativity. But things are not so simple! What the 'disturbance of particle 2' turns out to be is a *statistical* matter: the *probabilities* of outcomes of measurements on particle 2 are altered. And this kind of disturbance, it turns out, is not *necessarily* incompatible with Relativity. Whether the particular form of non-locality we have in quantum mechanics is or is not compatible with Relativity is searchingly examined in Maudlin ([1994]). One reason that 'non-local correlations' cannot be used to synchronize watches (which would immediately contradict Relativity) is that there does not seem to be any experimentally determinable fact of the matter as to when the 'signal' from Particle 1 (if there is a 'signal') *reaches* Particle 2—it could even reach Particle 2 *before* the measurement is made on Particle 1!

by Everett ([1957]) (although I subsequently discussed it in Putnam [1991] and Putnam and Albert [1995]). So there are many reasons for my returning to the subject and taking a second look at quantum mechanics.

As I did in the earlier paper, I will try to write in a way which is intelligible to readers who do not know quantum mechanics. I tried very hard to make the earlier paper self-contained and to explain *the logical structure* of the problem rather than to go into mathematical details, and that is what I hope to do again.

2 Scientific realism is the premise of my discussion[3]

I will begin by quoting Putnam ([1965], pp. 130–2), omitting footnotes and references. (I still agree with those pages, and they will set the stage for what I want to say now.)

> Before we say anything about quantum mechanics, let us take a look at the Newtonian (or 'classical') view of the physical universe. According to that view, nature consists of an enormous number of particles. When Newtonian physics is combined with the theory of the electromagnetic field, it becomes convenient to think of these particles as *dimensionless* (even if there is a kind of conceptual strain involved in trying to think of something as having a *mass* but not any *size*), and as possessing electrical properties—negative charge, or positive charge, or neutrality. This leads to the well-known 'solar system' atom—with the electrons whirling around the nucleus (neutrons and protons) just as the planets whirl around the sun. Out of atoms are built molecules; out of molecules, macroscopic objects scaling in size from dust motes to whole planets and stars. These latter also fall into larger groupings—solar systems and galaxies—but these larger structures differ from the ones previously mentioned in being held together exclusively by gravitational forces. At every level, however, one has trajectories (ultimately, that means the possibility of continuously tracing the movements of the elementary particles) and one has causality (ultimately that means the possibility of extrapolating from the history of the universe up to a given time to its whole sequence of future states).
>
> When we describe the world using the techniques of Newtonian physics, it goes without saying that we employ *laws*—and these laws are stated in terms of certain *magnitudes*, e.g. distance, charge, mass. According to one philosophy of physics—the operationalist view so popular in the 1930s—statements about these magnitudes are mere shorthand for statements about the results of measuring operations. Statements about distance, for example, are mere shorthand for statements about the results of manipulating foot rulers. I shall assume that this philosophy of physics is *false*.

[3] At the suggestion of Yemima Ben Menahem, let me say that by 'scientific realism' I mean what I called 'convergence realism' in Putnam ([1994a]). As I argued there, and in fact as far back as Putnam ([1976]), scientific realism in this sense does not presuppose either metaphysical realism or metaphysical antirealism, although it *is* incompatible with all forms of operationalism and phenomenalism.

> Since this is not a paper about operationalism, I shall not defend or discuss my 'assumption'. I shall simply state what I take the correct view to be.
>
> According to me, the correct view is that when the physicist talks about electrical charge, he is talking quite simply about a certain magnitude that we can distinguish from others partly by its 'formal' properties (e.g. it has both positive and negative values, whereas mass has only positive values), partly by the system of laws this magnitude obeys (as far as we can presently tell), and partly by its *effects*. All attempts to *literally* 'translate' statements about, say, electrical charge into statements about so-called observables (meter readings) have been dismal failures, and from Berkeley on, all *a priori* arguments designed to show that all statements about unobservables must ultimately reduce to statements about observables have contained gaping holes and outrageously false assumptions. It is quite true that we 'verify' statements about unobservable things by making suitable *observations*, but I maintain that without imposing a wholly untenable theory of meaning, one cannot even *begin* to go from this fact to the wildly erroneous conclusion that talk about unobservable things and theoretical magnitudes *means the same* as talk about observations and observables.
>
> Now then, it often happens in science that we make inferences from measurements to certain conclusions couched in the language of a physical theory. What is the nature of these inferences? The operationalist answer is that these inferences are *analytic*—that is, since, say 'electrical charge' *means by definition* what we get when we measure electrical charge, the step from the meter-readings to the theoretical statement ('the electrical charge is such-and-such') is a purely conventional matter. According to the non-operationalist view, this is a radical distortion. We know that this object (the meter) measures electrical charge *not* because we have adopted a 'convention', or a 'definition of electrical charge in terms of meter readings', but because we have accepted a body of theory that includes *a description of the meter itself in the language of the scientific theory*. And *it follows from the theory*, including this description, that the meter measures electrical charge (approximately, and under suitable circumstances). The operationalist view disagrees with the actual procedures of science by replacing a probabilistic inference within a theory by a non-probabilistic inference based on an unexplained linguistic stipulation.
>
> If the non-operationalist view is generally right (that is to say, correct for physical theory in general—not just for Newtonian mechanics), then *the term 'measurement' plays no fundamental role in physical theory as such*. Measurements are a subclass of physical interactions—no more or less than that. They are an important subclass, to be sure, and it is important to study them, to prove theorems about them, etc.; but 'measurement' can never be an *undefined* term in a satisfactory physical theory, and measurements can never obey any 'ultimate' laws other than the laws 'ultimately' obeyed by *all* physical interactions.

For myself, and for any other 'scientific realist', the whole so-called interpretation problem in connection with quantum mechanics is just this: *whether* we can understand quantum mechanics—no, let me be optimistic—*how* to understand quantum mechanics in a way that is compatible with the anti-operationalist philosophy that I subscribed to in the pages I just quoted, and that I have always subscribed to. But it took a long time for physicists to admit that there is such a problem. I can tell you a story about that. In 1962 I had a series of conversations with a world-famous physicist (whom I will not identify by name). At the beginning, he insisted, 'You philosophers just *think* there is a problem with understanding quantum mechanics. We physicists have known better from Bohr on.'[4] After I forget how many discussions, we were sitting in a bar in Cambridge, and he said to me, 'You're right. You've convinced me there is a problem here; it's a shame I can't take three months off and solve it.'

Fourteen years later, the same physicist and I were together at a conference for a few days, and he opened his lecture at that conference (a lecture which explained to a general audience the exciting new theories of quarks) by saying, 'There is no Copenhagen interpretation of quantum mechanics. Bohr brainwashed a generation of physicists.' Evidently, he had undergone a considerable change of outlook.

3 What 'quantum mechanics' *says*—and some problems

If I am to explain why there is a problem with understanding quantum mechanics in a way that is compatible with scientific realism, I first have to say what 'quantum mechanics' *says*. I will simplify considerably,[5] of course, but the logical structure of the problem will not be affected by these simplifications.

'Pure states' of a physical system—states about which we have as much information as quantum mechanics allows us to have—are represented by vectors in an abstract space called 'Hilbert space'. I will call these vectors 'state vectors' (and I will freely identify states and their state vectors to simplify exposition). Like vectors in a real vector space they can be multiplied by a scalar[6] and they are subject to appropriately defined operations of vector addition and inner product.

These states can also be represented as functions over what is called a 'basis'. For example, if we have a system of just two particles[7] A and B,

[4] The physicist in question was, of course, referring to Niels Bohr and the so-called Copenhagen interpretation of quantum mechanics.

[5] For example, I will not distinguish between a vector and a 'ray' (a one-dimensional subspace of a Hilbert. I will not employ 'bra-ket' notation, and, most importantly, I shall confine attention to pure states, ignoring the problem of understanding mixed states. On that problem, I still stand by what I wrote in Putnam ([1965], pp. 155–6).

[6] The scalars can be complex numbers, and not only real numbers, a fact which has a mathematical connection with the phenomenon of 'interference'.

[7] For the time being we confine attention to elementary non-relativistic quantum mechanics with a finite number of particles, and ignore spin.

and we let the position coordinates of A be x_1, x_2, x_3 and the position coordinates of B be x_4, x_5, x_6, then any pure state can be represented as a complex-valued function of x_1, x_2, x_3, x_4, x_5, x_6, t, where t represents *time*. This is called a 'representation in a *position basis*'.[8] (In general, if the system consists of N particles, the basis will have $3N$ coordinates.)

Naturally, we want to know how the states of a system change with time, and the answer is that, represented as functions in the way just described, they obey a famous differential equation, the Schrödinger equation. This equation is the heart of the 'dynamics' of quantum mechanics. (Whether this is the *only* way in which states can change is, as we will see, at the heart of the controversies about quantum mechanics.)

With just this minimum of information about what quantum mechanics says, I can already state the first problem that I want to discuss, the famous problem of 'Schrödinger's Cat'. I think everyone has heard of Schrödinger's Cat by now. Schrödinger imagined that one has a cat in a well-isolated system, say a satellite in space, and one has a device which shoots exactly one photon in the direction of a half-silvered mirror, beyond which is a detector. (The half-silvered mirror has the property that the quantum mechanical probability that the photon will go through is 1/2. If the photon goes through and hits the detector, in Schrödinger's bloody experiment the cat gets electrocuted;[9] if the photon is reflected, the cat lives.) According to the Schrödinger equation, the cat should end up in a 'state' which would be (represented by) the vector sum or 'superposition' of a vector which represents the cat surviving the experiment and a vector which represents the cat being electrocuted, a state of the form $\mathbf{1/\sqrt{2}}$(Live Cat) + $\mathbf{1/\sqrt{2}}$(Dead Cat). And Problem One is what are we to make of a state which is a superposition of two states like this, two states in which a *macro-observable* has different values? (Problem Two is simply that the existence of such a state is *unbelievable*, and this is connected with a joke of Einstein's that I will tell you later.)

If we never observe such a state, why don't we? All interpretations of quantum mechanics are required to give some answer to that question.

One version of the Copenhagen interpretation of quantum mechanics—a version which really *is* an interpretation, unlike Bohr's own remarks, which I interpret as a *rejection* of the possibility of a scientific realist interpretation—was given by Von Neumann ([1932]). According to Von Neumann's axiomatization, quantum mechanics says that physical states change in *two*

8 Instead of the position coordinates, x_1, x_2, x_3, x_4, x_5, x_6, we can also use the components of momentum in the corresponding directions as a basis, and there are still other possible choices. The function will be different, but the instructions for 'decoding' the function—interpreting it as the description of a state vector—will ensure that it represents the state we want it to.

9 In Putnam ([1994b]) and in Albert and Putnam ([1995]) the cat is tickled rather than electrocuted, the reader may be happy to know.

ways—and not always in accordance with the Schrödinger equation. When no measurement is made, changes do obey that equation. But when a measurement is made, states change in a different way: they 'collapse', according the Von Neumann axioms. This is so important that I will display it explicitly, and with more detail:

Von Neumann's interpretation

(i) When no measurement is made—and 'measurement' is a primitive notion; one is simply supposed to know what a 'measurement' is—states change in time in accordance with the Schrödinger equation.

(ii) When a measurement is made—when one measures the position of some system, or one measures its kinetic energy, or whatever—the state 'collapses' (changes discontinuously) into one in which the 'observable' (the physical quantity) which was measured has a definite value. Mathematically the collapse is represented not by a solution to a differential equation but by wiping out, erasing, the state that was there before and 'putting in by hand' the state that was found by the measurement. The probability that any given possible result of measurement (any given 'eigenstate' of the observable measured) will actually be the one found by the measurement is postulated to be proportional to the square of the absolute value of the vector which is the inner product of the unit vector corresponding to that eigenstate and the vector which we 'erased'—the vector which *would have been* the state of the system *if it had evolved according to the Schrödinger* equation.[10]

The answer to our Problem One—the problem of Schrödinger's Cat—on the Von Neumann interpretation is straightforward, but hardly satisfactory. We never see a cat in a state of the form $\mathbf{1}/\sqrt{\mathbf{2}}$(Live Cat) + $\mathbf{1}/\sqrt{\mathbf{2}}$(Dead Cat) because our looking to see whether the cat survived or not is a 'measurement', and so the cat obligingly collapses into the state (Live Cat) or into the state (Dead Cat) when we look!

4 Other interpretations of quantum mechanics

I mentioned at the start that today we know of many different interpretations of quantum mechanics. Table 1 presents two unconventional ones. The Bohm interpretation, first proposed in Bohm ([1952]), was rejected by me in Putnam

[10] An important terminological note: the word 'observable' in the foregoing has little or nothing to do with the epistemological use of 'observable' in my explanation of scientific realism at the outset of this essay. Unfortunately, 'observable' has become the standard term in quantum mechanics for any physical magnitude, whether we are able to measure it or not. It is not a philosophical term.

([1965]) for reasons I now want to retract, and the Ghirardi–Rimini–Weber (GRW) interpretation was proposed in Ghirardi et al. ([1986]), long after my paper was published.

Bohm's theory, a successor to the earlier 'pilot wave' interpretation described in De Broglie ([1960]),[11] has become the classical example of a hidden variable theory. According to it, particles do have definite positions and momenta at all times. They even have continuous trajectories. These trajectories are determined by two things: (i) a 'velocity field' and (ii) the initial positions and momenta of the particles, which are assumed to be distributed (at whatever is chosen to be the initial time t_0) randomly, but in accordance with the probability distribution given by the state vector. The 'velocity field' is determined by what is called 'the probability current' (which depends upon the state vector of the system in accordance with the classic rule for determining probabilities in quantum mechanics first suggested by Max Born, and employed by Von Neumann).

It is a consequence of the Bohm theory that these initial positions and momenta of the particles are in principle impossible to determine; all that we know about these initial positions is that they are distributed in such a way that the *probability* that any one of these particles is at any given place at t_0 is the standard quantum mechanical probability (and similarly for the momentum distribution). That property, that the probability distribution is quantum mechanical, is then preserved through all time. In this way, the Bohm interpretation explains why the probability of finding any given particle in any given place, or, indeed, of finding the whole system of particles in any given position configuration at any given time, is in accordance with standard quantum mechanical calculations (e.g. with those given by Von Neumann).

In Putnam ([1965]), I rejected Bohm's interpretation for several reasons which no longer seem good to me. Even today, if you look at the Wikipedia encyclopaedia on the Web, you will find it said that Bohm's theory is mathematically inelegant. Happily, I did not give *that* reason in Putnam ([1965]), but in any case it is not true. The formula for the velocity field is extremely simple: you have the probability current in the theory anyway, and you take the velocity vector to be proportional to the current. There is nothing particularly inelegant about that; if anything, it is remarkably elegant! However, I rejected Bohm's theory for reasons I got from my teacher, Hans Reichenbach. In Reichenbach ([1944]), Bohm's theory is rejected as leading to unacceptable 'causal anomalies'. I was also misled by Bohm's unfortunate choice of a term for the velocity field; he called it a 'potential'.

[11] The idea of interpreting the state functions of quantum mechanics as 'pilot waves' was also entertained early on by Schrödinger, who seems to have quickly given it up.

Table 1 Two unconventional interpretations of quantum mechanics

Bohm	*GRW (Ghirardi–Rimini–Weber)*
The classical hidden variable theory	The continuous spontaneous localization model
Particles do have definite positions and momenta all the time. The trajectories are determined by a velocity field which is itself determined by the quantum mechanical 'state'	Each particle has a tiny probability of spontaneously jumping into a definite position state. If there are many millions of atoms, an object will, as a result, always have a definite position (thus solving the 'problem of Einstein's bed')

I took it to be literally a potential energy distribution and argued that that is an interpretation which it will just not bear.

It is certainly true that the Bohm theory implies certain 'causal anomalies'. But the most obvious of these are no longer reasons for rejecting the theory, because we know that they really occur in nature. Bohm's theory implies non-locality. Both Reichenbach and I worried about the question 'How come this Bohm "potential" (which we thought of as a force) doesn't get weaker with distance?' The answer, as we now know, is 'Because non-local correlations can appear over any distance.' If, as Maudlin ([1994]) has suggested, we think of the 'Bohm field' as *a mathematical representative of non-locality*, then we need no longer be bothered by the 'causal anomaly'. In sum, I do not think the Bohm theory can simply be dismissed, removed from consideration, in the way I did in 1965.

In the interpretation on the right of Table 1, the interpretation proposed by Ghirardi et al. ([1986]), which is called 'the spontaneous continuous localization model', each particle has a *tiny* probability—*truly* tiny—of spontaneously jumping into a definite position state. The probability is so tiny that, for example, if the system consists of just one isolated hydrogen atom, one would have to wait, on average, many thousands of years for it to jump into a definite position state. In other words, don't hold your breath. *But*, as we all know, an object such as the table in front of me as I type these words consists of millions upon millions of particles. If there are many millions of particles, the object will, according to GRW theory, always have a definite position.[12] As I say in Table 1, that solves 'the problem of Einstein's bed', and I will tell you what that problem is in a moment.

12 More precisely, it will have a definite position by macroscopic standards of definiteness with a probability which is so close to one that it can be treated as certainty.

The reason it solves the problem is that, according to quantum mechanics, if even one of the particles, even one of the electrons, for example, of which the table consists, 'jumps' into a definite position state, that state *multiplies* the state of the whole system by a factor which forces that state to become definite with respect to position (to come close to being an eigenstate of position). These 'spontaneous collapses' may be happening at times which are very far apart, as far as any one particle is concerned, but they are happening *all the time* as far as this very large collection of particles is concerned. For that reason, the probability that the table will fail to have a definite position (by macroscopic standards of definiteness) becomes virtually zero. It is not *impossible*, but it will never happen.

5 The problem of Einstein's bed

Why do I call this 'the problem of Einstein's bed'? I said early on that the existence of superpositions of states in which macro-observables have different values, states such as $\mathbf{1/\sqrt{2}}$(Live Cat) + $\mathbf{1/\sqrt{2}}$(Dead Cat), is *unbelievable*, and that the problem this poses is connected with a joke of Einstein's. Here is the story of Einstein's joke.

I met Einstein just once, in 1953. He was a friend of my teacher, Hans Reichenbach. When I joined the Princeton faculty in 1953, Reichenbach arranged for me to meet Einstein. I had tea with him in the little house on Mercer Street. Not unexpectedly, he talked about his dissatisfaction with quantum mechanics. He did *not* say, 'God doesn't play with dice.' He did not say or imply anything to the effect that he could not accept a theory which is indeterministic (and we now know from Howard [1985]) and other historians who have studied Einstein's unpublished correspondence that the failure of determinism was not Einstein's most significant problem with quantum mechanics). What he said on that occasion was something like the following: 'Look, I don't believe that when I am not in my bedroom my bed spreads out all over the room, and whenever I open the door and come in it jumps into the corner.'

In other words, Einstein could not believe Von Neumann's 'collapse' assumption.

To tell you the truth, I do not like the term 'measurement problem', which is often used in connection with quantum mechanics. I think 'collapse problem' would be a better term. In other words, the real problem is *Do we or don't we need to postulate a 'collapse', and if we do assume a 'collapse', what should we say about it?* Ever since GRW has been in the field, there have been interpretations of quantum mechanics in which collapse has nothing to do with *measurement*. It happens whether a human makes a measurement or not. For that reason, the term 'measurement problem' is out of date.

In Putnam ([1965], pp. 149–55), I said that the view that I regarded as the most plausible one to take was a variant of the Copenhagen interpretation.[13] In order to avoid taking 'measurement' as a primitive notion, I said that Von Neumann's assumption that the state 'collapses' into an eigenstate of the observable measured upon measurement should be replaced by a different assumption

> Instead of saying henceforth that observables do not exist unless measured, we will have to say that *micro*-observables do not exist unless measured. We will take it as an assumption of quantum mechanics that *macro*-observables retain sharp values (by macroscopic standards of sharpness) at all times. The formulation of the Copenhagen interpretation that I am now considering, then, comes down to this: that macro-observables have sharp values at all times in the sense just explained, while micro-observables have sharp values only when measured, where measurement is to be defined as a certain kind of interaction between a micro-observable and a macro-observable.

And I said (Putnam [1965], p. 157) that the remaining open problem for quantum mechanics was to say what is so special about macro-observables: 'The result we wish is that although micro-observables do not necessarily have definite numerical values at all times, macro-observables do. And we want this result to come out of quantum mechanics in a natural way. We do not want simply to add it to quantum mechanics as an *ad hoc* principle. So far, however, attempts to derive this result have been entirely unsuccessful.'

6 Classification of the possible kinds of interpretation

For many years after that, I tried to come up with an interpretation of quantum mechanics that would solve the problems I have described. In Putnam ([1968]), I even proposed an interpretation which involved a non-standard logic, but, as I explained in Putnam ([1994b]), that interpretation collapsed (though not in the quantum mechanical sense of 'collapse'). More recently, however, it has occurred to me that I should instead attempt to classify *all possible interpretations* (or all possible interpretations that do not involve giving up classical logic, since I had satisfied myself that the approach via a non-standard logic would not work). In this way, I arrived at the four classes shown in Table 2. That these are the possibilities is, then, the first fruit of my investigation, and obviously it rests on little more than the law of the excluded middle: either an interpretation says there is a collapse or it does not; if it says there is a collapse then either that collapse is spontaneous

[13] Here I understood 'Copenhagen interpretation' to mean Von Neumann's axiomatization, not Bohr's rejection of scientific realism.

Table 2 Kinds of interpretations of quantum mechanics

Collapse	*No collapse*
Produced by something external to the system and not subject to superposition (e.g. Von Neumann)	No hidden variables (Many Worlds)
versus	versus
Spontaneous (e.g. GRW)	Hidden variables (e.g. Bohm)

or it is explained by something external to the system (that does not itself undergo collapse—otherwise we get a regress of collapses); if it says there is no collapse, then either there are hidden variables or there are not. After the possible interpretations are laid out, each of us can say which of these possibilities should be ruled out as a 'non-starter'—as I will now attempt to do—but it seems to me that considerable clarification will result if we can agree that these *are* the possibilities before us if we want a scientific realist interpretation.

In my table of the four kinds of possible (scientific realist) interpretations, I give an example of the currently most plausible version of each, to stand for the kind, but really I am interested in these as *classes* of possible interpretations.

On the left we have collapse interpretations and on the right we have the no-collapse interpretations. (Even Brouwer would have to agree that this is a legitimate use of the law of the excluded middle!) Either an interpretation of quantum mechanics says that there is a 'collapse of the state function' or it says that there is no collapse, that it never happens.

In the left column, the 'collapse' column, the first possibility is that the collapse is produced by something external to the system and not subject to superposition, which is Von Neumann's proposal. For example, one might propose that macro-observables are not subject to superposition, as I did in my 1965 paper, or one might say—Von Neumann hints at this in his book, and Eugene Wigner famously advocated it—'No, the collapse occurs when the result of a measurement is *registered by a consciousness*.' I do not know of anyone who currently advocates this 'psychical' view, but those are the two 'classical' versions of the view that collapse is produced by a 'measurement', where 'measurement' means that the system interacts with something that is intrinsically not subject to superposition.

However—as GRW has taught us—there is another possibility to be included in the 'collapse' column: the collapse could be spontaneous. For example, it could be an ultimate statistical law of nature that a particle has a certain fixed probability of 'jumping into a position eigenstate' (or, as in GRW, a state very close to but not exactly the same as a position eigenstate). Alternatively, the spontaneous collapse could be provoked by a 'trigger'—it could, for example, as suggested by Penrose ([1994]), occur whenever the global gravitational field would otherwise be superimposed to a certain unacceptable extent, although this 'gravitational trigger' suggestion has not, to my knowledge, been worked out in detail.

What about the interpretations in the other column—the 'no-collapse' column?

The most famous of these is the so-called Many Worlds interpretation. This interpretation was proposed by the late Hugh Everett III, who suggested in 1957, in a paper which, as Albert ([1992], pp. 112–113) has put it, is 'both extraordinarily suggestive and hard to understand', that he had found a way of coherently entertaining the possibility that there is no collapse and there are no hidden variables: the world (the state function of the whole universe) just rolls on and on as predicted by the Schrödinger equation. Again quoting Albert, 'The idea of what has become the canonical *interpretation* of Everett's paper Albert has in mind (Dewitt [1970]) is that the means of coherently entertaining the possibility that Everett must have had in mind (or perhaps the one that he *ought* to have had in mind) is to take the two components of a state [such as $\mathbf{1}/\sqrt{\mathbf{2}}$(Live Cat) + $\mathbf{1}/\sqrt{\mathbf{2}}$(Dead Cat)][14] ... to represent (literally!) two physical *worlds*.' The idea, applied to the thought experiment of Schrödinger's Cat, is that when the photon reaches the half-silvered mirror the number of physical worlds there are literally increases from one to two: in one of the two worlds the photon goes through the half-silvered mirror, the cat is electrocuted, and the observer has the determinate belief that the cat is dead; in the other the photon is reflected, the cat lives, and the observer (and presumably the cat as well) has the determinate belief that the cat is alive.

Finally, there are no-collapse interpretations in which there *are* hidden variables, and the most famous of these is the Bohm interpretation that I described earlier. In that interpretation, positions are the hidden variables: they can sometimes be localized, it is true, but there is no way of determining the precise trajectories that the particles follow.

(But why did I say there is no collapse in the Bohm interpretation? What happens to the state function when I make a position measurement on a particle, and I say 'now it is here'? The answer is that the state continues to evolve according to the Schrödinger equation on the Bohm interpretation, exactly as in the Many Worlds interpretation! Bell, who showed that quantum mechanics implies non-locality, referred to the particle locations in the

[14] Albert uses a different example.

Bohm interpretation—which he favoured—as the 'be-ables', rather than as 'hidden variables'. These be-ables are physical quantities which have definite values at all times. The quantum mechanical state 'guides' the particles, but it never collapses, on the Bohm interpretation.)

So these are the four families of interpretations. To repeat: either there is a collapse or there is no collapse. If there is a collapse, either it is produced by something not subject to superposition, as in Von Neumann's formulation of quantum mechanics, or it is spontaneous. If there is no collapse, either there is no collapse and no hidden variable, no be-ables, or there is no collapse plus be-ables. That is how I propose we draw the map. Even if we subsequently rule out some of the interpretations we should not *begin* by ignoring them. Now I will say which interpretations in Table 2 I think we should discard.

7 Which interpretations I think we can rule out

As I have said, in Putnam ([1965]) I chose the left column ('collapse'), and I chose the first interpretation in that column, taking macro-observables to be the 'something external to the system and not subject to superposition' that brings about the collapse, and I took it to be an unsolved problem why this is the case. I took it to be a problem because it cannot be an ultimate postulate of physics that macro-observables do not undergo superposition. 'Macro-observable' is not the sort of term that can be an irreducible primitive in an ultimate physical theory, so I called for some future extension of quantum mechanics that would explain *why* macro-observables do not go into such states as $\mathbf{1}/\sqrt{\mathbf{2}}$(Live Cat) + $\mathbf{1}/\sqrt{\mathbf{2}}$(Dead Cat).

That proposal seems less and less attractive to me. That there is something special about macro-observables seems tremendously unlikely, unless what is special is, for example, that they involve large numbers of particles and the particles spontaneously collapse, or something of that sort. So I am now inclined to give up this most *classic* interpretation—it was, in a way, *the* interpretation that was actually used in the first decades of quantum mechanics. And this leaves the second interpretation—spontaneous collapse—as the open possibility in the left hand column.

I do not mean necessarily GRW theory. There are certain problems with the GRW account of collapse. For example, if a system collapses into an eigenstate of position at a certain time—if it collapses into a really *absolutely sharp* position state (which Ghirardi, Rimini, and Weber do not say, for just the reason I am going to describe)—then, according to quantum mechanics, the 'complementary' observable, momentum, would be infinitely indeterminate. It could be as huge as you like, and then one would have violations of the conservation of energy. So, what GRW postulates is that the collapse is not infinitely sharp, but also not too 'spread out' (otherwise Einstein's bed would lose its nice definite position). The upshot is that GRW predicts violations of the conservation of energy that are *too small to be observed*. This is a little distressing (but then life is hard, you can't please everyone!). Anyway, if GRW turns out to be true, I hereby predict that somebody will come up with a redefinition of 'energy' which saves the law of the conservation of energy. (Mathematicians are very clever, and the history of physics suggests that they would somehow manage.)

On the other side of my table, the 'no collapse' side, the Many Worlds interpretation now has distinguished advocates (e.g. Gell-Mann and Hartle [1990]). I myself think it is untenable, and my reason is very simple.

Suppose I perform an experiment. It can have just two outcomes, as in the Schrödinger's Cat thought experiment, except that instead of having a complicated system such as a cat, let us just have two light bulbs, a red one and a green one, and let the experimental set-up be arranged so that either the red light goes on or the green light goes on when the interaction takes place (but not both). And the set-up need not be such that the two outcomes are equiprobable, as was the case with Schrödinger's Cat. The two outcomes could have probabilities of 0.9 versus 0.1, or 0.00001 versus 0.99999, or whatever you wish. The probabilities can be very unequal, and it does not matter.

I perform this experiment. Then I perform it again. Then I perform it a third time. Let us suppose I repeat it thirty times.

According to the Many Worlds interpretation, since the world obeys the Schrödinger equation at all times, in the course of the first trial of the experiment it really does go into a superposition of the form $\boldsymbol{p}$(Green Light lights) + $\boldsymbol{q}$(Red Light lights), where both $\boldsymbol{p}$ and $\boldsymbol{q}$ have positive absolute values, and $(\mathbf{1} - \boldsymbol{p})$ is the probability that the red light goes on, and *this* means, on Everett's interpretation as interpreted by Dewitt, that there are now (after the first trial) *two* physical universes, in one of which I observe the green light go on (and only that light go on) and in the second of which I observe only the red light go on.

So far this is just a simplified version of 'Einstein's bed'. If Einstein had chosen to talk about this experiment, rather than (unfortunately for his own case[15]) about his bed, he would have asked, 'Why don't I see a *superposition* of red light on and green light on?' And the answer that Everett–Dewitt would have given him would have been 'Because the state vector $\boldsymbol{p}$(Green Light lights) + $\boldsymbol{q}$(Red Light lights) is *coupled* with Einstein's state vector,

[15] I say 'unfortunately for his own case' because the Schrödinger equation does *not* imply that the bed would go into a superposition of states in which it is in different places in the room between times when Einstein enters the bedroom.

the Einstein in the *first* of the two physical universes that exist after the interaction *sees* the green light go on, and the Einstein in the *second* of the two physical universes *sees* the red light go on.' In the total reality—we might call it the Everett 'multiverse'—after the experiment is performed the first time, there will be two Einsteins, one seeing a green light shining and the other seeing a red light shining. After the experiment is performed a second time, there will be *four* Einsteins, and four Einstein-histories: (i) a history which goes Einstein sees green followed by green again, (ii) a history which goes Einstein sees green followed by red, (iii) a history which goes Einstein sees red followed by green, and (iv) a history which goes Einstein sees red followed by red. After the experiment has been performed thirty times there will be 2^{30} Einsteins, with 2^{30} histories. And, unlike Schrödinger's Cat, which is only a thought experiment, performing a simple two-outcome experiment 30 times is something any of us can do.

I repeat, on the Many Worlds interpretation, there will be 2^{30} Einstein-histories—'parallel worlds'; science fiction is literally right!

Now, suppose that instead of saying 'Einstein', I say 'I' (and I ask you, dear reader, to think of yourself as the 'I' in question). Suppose *I* do this. And I ask myself, '*What is the probability in the naïve sense*—not the "probability" in the quantum mechanical sense, this real number which I calculate by finding the square of the absolute value of a certain vector, but the probability in the sense of *the number of my future histories in which I will observe that, say, the green light went on half of the time plus or minus 5% of the time divided by the total number of my future histories*?' You can figure this out by simple combinatorics. (All it takes is the binomial theorem.) And the answer is *independent of the quantum mechanical 'probability'*. So why should I use quantum mechanical 'probability' to predict what I am likely to observe?

Not surprisingly, the largest number of my future histories are ones in which the green light will have gone on approximately 50% of the time, and this is so *regardless of what the quantum mechanical 'probability' happens to be*.

Let me put it another way: if histories in which what one of the 2^{30} Hilary Putnams observes *confirms* quantum mechanics are so rare—as they will be if $\boldsymbol{p}$ is very small, as we can easily arrange for it to be—*how come we are so lucky*? How come *my* observations have so far been in accordance with quantum mechanical probability? On the Many Worlds interpretation, quantum mechanics is the first physical theory to predict that *the observations of most observers will disconfirm the theory*.

The philosophical moral behind my question is this: once you give up the distinction between actuality and possibility—as the Many Worlds interpretation in effect does, by postulating that all the quantum mechanical possibilities are actualized, each in its own physical universe—once you say that all possible outcomes are, ontologically speaking, equally actual—the notion of 'probability' loses all meaning. 'No collapse and no hidden variables' is incoherent.[16]

8 The 'moral' of this discussion

What we are left with, if what I have said so far is right, is a conclusion that I initially found very distressing: *either GRW or some successor, or else Bohm or some successor, is the correct interpretation*—or, to include a third possibility to please Itamar Pitowski,[17] we will just fail to find a scientific realist interpretation which is acceptable. (And the ghost of Bohr will laugh, and say, 'I told you all along that the human mind cannot produce a realist interpretation of quantum mechanics'!) But if we are optimists and think that there is somehow a realistic interpretation to be found, then—as argued in Maudlin ([1994])—we are left with GRW and Bohm.

Now why did this conclusion make me unhappy? Not just because of the paradoxical consequences of each of these theories (after all, the phenomena they are called upon to account for are 'paradoxical'). In the case of the Bohm theory, there is the fact that it has not yet been extended to a quantum field theory, let alone to a quantum cosmology, although Goldstein and Teufel ([2001]) have already made proposals in this direction. It is not just the technical problems—which are a good thing, a stimulus to productive research. It is that neither of these theories is Lorentz invariant, and it seems likely to me there is no rigorous proof, but, as Maudlin argues, it is pretty clear that no theory in either of the classes that they represent (the 'no collapse and hidden variables class' and the 'spontaneous collapse' class) can do without an 'absolute time' parameter. An absolute time will come back into the picture if either *sort* of theory is destined to be the future physics.[18]

[16] Defenders of the Many Worlds interpretation often appeal to a set of 'decoherence' theorems to show why detectable interference between the different 'histories' does not occur. These theorems answer objections to the interpretation from *within* the Many Worlds picture. But if my argument above is correct, the entire picture is incoherent, and the decoherence theorems in no way speak to that problem.

[17] Itamar Pitowski was my very helpful and challenging commentator at the meeting of the Israel Association for History and Philosophy of Science at which this article was delivered.

[18] I do not mean to imply that the existence of an absolute time parameter is, by itself, incompatible with Lorentz invariance; David Albert informs me that his student Catherine Peters has given a nice example of a space-time (a 'cylindrical' space-time) which has a Minkowski metric, and is thus locally Lorentz invariant, but in which only one rest system possesses well-defined global simultaneity surfaces. But (at present) we have no reason to think that this purely mathematical possibility has physical significance.

(My commentator at the conference at which this paper was read, Itamar Pitowski, said, 'You are saying that before we can interpret quantum mechanics we have to *change* it.' My reply was that Von Neumann *already* 'changed' quantum mechanics, certainly from Bohr's point of view. *All* interpretations of quantum mechanics are in a sense 'changes' of quantum mechanics, because it is an *incomplete theory*—one cannot 'regiment' it, formalize it in standard logical notation (e.g. quantificational logical notation) *unless* you add an 'interpretation'. And up to now, Putnam's law, the law that an interpretation of quantum mechanics is believed by its inventor and up to six other physicists, has held good.)

Can we soften the 'bad news' that we may need to return to a notion of 'absolute time'? My final suggestion is this: when it comes to quantum cosmology—and, as yet, neither GRW nor the Bohm theory has been extended to quantum cosmology—in my view, *present-day* quantum cosmology *does already* involve a 'background' time parameter. It is sometimes concealed, as when cosmologists say that they are not really taking an absolute time as the parameter in the Schrödinger equation but are taking something such as the 'radius' of the universe as the time parameter (and hoping that this is a well-behaved quantity). But this parameter plays exactly the role of an absolute time in which the cosmos is supposed to evolve.

The reason for its presence is that, in present-day quantum cosmology, one does not talk about *one single space-time*. Quantum mechanics depends on the idea that all physical 'states of the world' (Live Cat, Dead Cat, Red Light lights, Green Light lights, and so on) are represented by mathematical objects, vectors, which can be multiplied by scalars and can be added, so that one gets such states as $\boldsymbol{p}$(Green Light lights) + $\boldsymbol{q}$(Red Light lights). In quantum cosmology, the state vectors can represent *different geometries of space-time*. (A classic presentation is Misner et al. [1973].) In effect, one superimposes whole space-times. And this superposition of space-times evolves in the background time.

So, what relieves my initial distress at the idea of an absolute time coming back into the picture is the following thought: it might not be quite as bad a contradiction of Einstein's vision as it first seems. It might be that, before we 'superimpose', *each space-time is perfectly Einsteinian*—each space-time is a Minkowski space-time which knows nothing about any 'simultaneity'. And it may be that the time parameter that both GRW and Bohm need is just the absolute time parameter that quantum cosmology seems to need. Of course, this is just a speculation. But it would mean that, although Einstein would have to admit that there is such a thing as simultaneity, it comes from 'outside' any one well-defined space-time, it comes from the quantum mechanical 'interference' between whole space-times. And with this speculative suggestion, I will close.

Acknowledgements

The views I defend in this paper owe a great deal to Maudlin ([1994]) and to David Albert, with whom I have discussed these issues for more than twenty years.

Department of Philosophy
Harvard University
Cambridge, MA 02138
USA
hputnam@fas.harvard.edu

References

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

Albert, D. and Putnam, H. [1995]: 'Further Adventures of Wigner's Friend', *Topoi*, **14(1)**, pp. 17–22.

Aspect, A., Dalibard, J. and Roger, G. [1982]: 'Experimental Test of Bell's Inequalities Using Time Varying Analyzers', *Physical Review Letters*, **49**, pp. 1804–7.

Bell, J. [1964]: 'On the Einstein–Podolsky–Rosen Paradox', in J. Bell, 1987, *Speakable and Unspeakable in Quantum Mechanics*, Cambridge: Cambridge University Press.

Bohm, D. [1952]: 'A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables: I and II', *Physical Review*, **85**, pp. 166–93.

De Broglie, L. [1960]: *Non-linear Wave Mechanics: A Causal Interpretation*, Amsterdam: Elsevier.

Dewitt, B. [1970]: 'Quantum Mechanics and Reality', *Physics Today*, **23**, pp. 30–5.

Everett, H., III [1957]: 'Relative State Formulation of Quantum Mechanics', *Reviews of Modern Physics*, **29**, pp. 454–62. Reprinted in Wheeler, J. and Zurek W. (*eds*.), 1983, *Quantum Theory and Measurement*, Princeton, NJ: Princeton University Press.

Finkelstein, D. [1987]: 'The Quantum Paradox', in *Encyclopedia Britannica Yearbook of Science and the Future*, pp. 186–207.

Gell-Mann, M. and Hartle, J. B. [1990]: 'Quantum Mechanics in the Light of Quantum Cosmology', in W. Zurek (*ed*.), *Complexity, Entropy and the Physics of Information*, vol. 8, Santa Fe Institute Studies in the Sciences of Complexity, Boston: Addison-Wesley.

Ghirardi, G. C., Rimini, A. and Weber, T. [1986]: 'An Attempt at a Unified Description of Microscopic and Macroscopic Systems', in V. Gorini and A. Frigerio (*eds*.), *Fundamental Aspects of Quantum Theory*: *Proceedings of the NATO Advanced Research Workshop, Como, September 2–7, 1985*, NATO ASI Series B 144, New York: Plenum Press, pp. 57–64.

Goldstein, S. and Teufel, S. [2001]: 'Quantum Spacetime Without Observers: Ontological Clarity and the Conceptual Foundations of Quantum Gravity', in Callender, C. and Huggett, N. (*eds*.), *Physics Meet Philosophy at the Planck*

Scale; Contemporary Theories in Quantum Gravity, Cambridge: Cambridge University Press, pp. 275–89.

Howard, D. [1985]: 'Einstein on Locality and Separability', *Studies in History and Philosophy of Science*, **16**, pp. 171–201.

Maudlin, T. [1994]: *Quantum Non-locality and Relativity: Metaphysical Intimations of Modern Physics*, Oxford: Basil Blackwell.

Misner, C. W., Thorne, K. S. and Wheeler, J. A. [1973]: *Gravitation*, San Francisco: W.H. Freeman.

Penrose, R. [1994]: *Shadows of the Mind*, Oxford: Oxford University Press.

Putnam, H. [1965]: 'A Philosopher Looks at Quantum Mechanics', in H. Putnam, 1975, *Philosophical Papers*, vol. 1, *Mathematics, Matter and Method*, Cambridge: Cambridge University Press, pp. 130–58.

Putnam, H. [1968]: 'The Logic of Quantum Mechanics', in H. Putnam, 1975, *Philosophical Papers*, vol. 1, *Mathematics, Matter and Method*, Cambridge: Cambridge University Press, pp. 174–97.

Putnam, H. [1976]: 'Realism and Reason' (Presidential Address to the Eastern Division of the American Philosophical Association), in H. Putnam, 1978, *Meaning and the Moral Sciences*, London: Routledge and Kegan Paul.

Putnam, H. [1991]: 'The Uncertainty Principle and Scientific Progress', *Iride*, **7**, pp. 9–27.

Putnam, H. [1994a]: 'Three Kinds of Scientific Realism', in H. Putnam, *Words and Life*, Cambridge, MA: Harvard University Press.

Putnam, H. [1994b]: 'Michael Redhead on Quantum Logic', in P. Clark and B. Hale (*eds*.), *Reading Putnam*, Oxford: Blackwell, pp. 265–80.

Reichenbach, H. [1944]: *Philosophic Foundations of Quantum Mechanics*, Berkeley: University of California Press.

Von Neumann, J. [1932]: *The Mathematical Foundations of Quantum Mechanics*, transl. R. T. Beyer, 1955, Princeton, NJ: Princeton University Press.