CHAPTER 15

DISTILLING METAPHYSICS FROM QUANTUM PHYSICS

TIM MAUDLIN

METAPHYSICS is the theory of being, that is, the most generic account of what there is. As such, it must be informed by empirical science, since we can only discover the nature of the material world through our experience of it. The most general and fundamental account of material reality is provided by physics, hence physics is the scientific discipline most closely allied to (if not continuous with) metaphysics as a philosophical inquiry.

Modern physics has been an especially fertile source for astonishing suggestions about reality. No respectable inquiry into the nature of space and time, for example, can afford to ignore the Theory of Relativity, whose account of spatio-temporal structure would not have been discovered by any amount of armchair reflection. As truth is stranger than fiction, so is actual physical theory more conceptually challenging than a priori speculation.

This is not to say that as philosophers we should trade in our tools for those of the physicists. Physics provides theories which typically consist of a mathematical formalism and some procedures for applying that formalism to particular concrete situations. But both the formalism and the procedures may admit of alternative ontological *interpretations*. It may not be clear, for example, which part of the mathematics corresponds to real physical magnitudes and which is an artefact of arbitrary choices of units or gauges. It may not be clear which mathematical models represent real physical possibilities, and which do not. And it may not be clear which pairs of mathematical models represent the same physical situation. All of these problems confront even the philosopher who tries to take, for example, the Theory of Relativity 'at face value'.

These problems are magnified exponentially when one seeks to understand the ontological implications of quantum theory. There one finds a mathematical formalism and a set of practical procedures for using it, but no uniformity of opinion about bow that formalism is to be interpreted. Further, there is almost nothing about which the alternative available interpretations agree, or which can be directly inferred from even the most surprising experimental phenomena.

If the ontological ramifications of quantum theory depend so critically on how the formalism is interpreted (or, in some details, on exactly what the formalism is), then tbe right way to address our topic would be to present the various interpretations in detail. Unfortunately, limitations of space make this impossible. A beautiful account of the interpretations can be found in David Albert's *Quantum Mechanics and Experience* (Albert 1992), and any philosopher with a serious interest in the subject should peruse that book. Instead of a proper explication, then, I shall here present only the merest sketch of the interpretations, along with a running tally of what each would imply about various metaphysical claims which have been made about quantum theory.

In order to be a version or interpretation of quantum theory at all, a theory must make use of a *quantum state* or *wavefunction*. If one begins with a classical theory, there are rules of thumb about how to produce a corresponding quantum theory, with different sorts of classical systems yielding different sorts of quantum states. Roughly speaking, a classical system of a given sort can, at any moment, be in one of a number of possihle *configurations*: the configuration a system of particles is given by specifying the location of each particle; the configuration of a classical field is given by specifying the field's value at every point in space. If we think of each distinct configuration as a point in an abstract space, then the collection of all possible configurations forms the *configuration space* of the system.[1] A single point in configuration space specifies the complete instantaneous state of a classical system, and a trajectory through configuration space (parameterized by time) represents a complete physical history of the system.[2] The wavefunction of a system is typically a complex-valued function *on the configuration space*, i.e. a function which assigns a complex number to each possible configuration.

[1] There may be disputes about what ought to count as distinct configurations of a system (e.g. if all the particles in a system are moved by some fixed amount, is that a distinct configuration?), but I will ignore any such problems here.

[2] A configuration is the *instantaneous* state of a system, and so does not specify, for example, the velocities of particles, but only their positions. The velocities can of course be calculated if one is given a time-parameterized trajectory through configuration space.

If the system consists of a single classical particle, then a configuration of the system is just a specification of where the particle is, so the configuration space of the system is isomorphic to physical space (the collection of all the possible configurations of the system is given by the collection of all spatial locations, which is physical space). *In this case*, it is easy to fall into picturing the wavefunction of the particle as a kind of classical field, spread out in space, especially when thinking of, for example, the notorious two-slit experiment. But it can be severely misleading to assimilate wavefunctions in general to classical fields. If the system consists of two particles, then each point in the configuration space specifies the location of *both* the particles. One cannot, in this case, sensibly ask what the value of the wavefunction is *here* (indicating a point in physical space); one must ratber ask what tbe value of the wavefunction is for the configuration in which one particle is here and the other at some other particular location. This feature of the wavefunction will be of paramount importance when we come to the issue of *entanglement*, but most of the issues I will examine can be illustrated using only a single particle as the system, and I will use this simple system wherever possihle.

Although there are decent rules of thumb for quantizing many classical theories, two caveats should be kept in mind. First, not every classical theory has a straightforward quantum version, the General Theory of Relativity being the prime example. Secondly, not every quantum theory need be the quantized version of a classical theory: string theory, in particular, has been claimed to be a pure quantum theory, not obtainable by quantizing any classical system.

Incidentally, quantizing a theory, in this sense, means specifying a set of quantum states and associating certain mathematical operators with the classical quantities; it does not necessarily mean making classically continuous quantities discrete. It is true that certain quantized systems (e.g. the simple harmonic oscillator) have only a discrete set of allowable values for some quantities (e.g. energy) which can take a continuum of values in classical physics. But the position of a free quantum-mechanical particle is not 'quantized' in this sense: it can take any of a continuum of values.

All 'interpretations' of quantum theory, then, employ a quantum state or wavefunction. They all also agree on *part* of the dynamics for that wavefunction: the wavefunction at least usually develops in accord with a particular sort of deterministic linear equation of motion. The *Schrödinger* equation is used in the quantum theory of non-relativistic particles.[3] Again, if only a single particle is under consideration, the wavefunction can be visualized as a field in pbysical space, and tbe Schrödinger equation then specifies how that field changes with time. This gives rise to the usual (and somewhat misleading) picture of an electron 'smeared out in space' passing through both slits in the two-slit experiment and then (somehow)

[3] One uses a different equation, the Dirac equation, for Special Relativistic contexts, and there are other equations for field theories of various sorts.

giving rise to interference bands on the detection screen. But it is not that anything is smeared out in space; rather the wavefunction 'propagates' in configuration space in accord with the Schrödinger equation. All interpretations agree, though, that the state of the wavefunction depends on whether each slit is open or closed, and that this dependence on both slits plays a central role in accounting for the interference. There the agreement among interpretations ends.

The problem with positing *only* the wavefunction governed by Schrödinger evolution was pointed out long ago by Schrödinger himself, in his notorious example of the cat. A variation of the problem can be posed with polarized light. Suppose we have a polarized filter oriented in the vertical direction with a detector placed behind it. If we send a vertically polarized photon towards the filter, it will pass and the detector will fire. If we send a horizontally polarized photon, it will not pass and the detector will not fire. These plain physical facts are reflected in the behaviour of the wavefunction: starting with the quantum state of the vertically polarized photon and the detector in its ready state, Schrödinger evolution of the wavefunction yields a final state in which the only non-zero part of the wavefunction corresponds to the configuration in which the detector has fired. Similarly, the wavefunction of a horizontally polarized photon and the detector will evolve so the only non-zero part of the wavefunction corresponds to the detector not firing. But what if we send a *diagonally* polarized photon at the filter, as we can easily arrange?

The quantum state of a diagonally polarized photon is, as a mathematical object, a vector sum of the state of vertical polarization and the state of horizontal polarization. The *linearity* of the Schrödinger equation then implies that the end state of the wavefunction, after the interaction of the photon with the screen and the detector, will be a vector sum of a quantum state in which the detector has fired and a state in which it has not. The wavefunction in this case has a non-zero value for the detector firing and a non-zero value for the detector not firing, and those values may be made equal. The end state is, as we say, a superposition of a state in which the detector fires and one in which it does not. The central interpretative question of quantum theory is what we are to make of this state.

As a practical matter, physicists discovered long ago how to use this state for making predictions. Since, it seems, the actual detector either fires or fails to fire on any particular run, the final wavefunction should be used to assign *probabilities* to these two outcomes. The method for doing this is simple and uncontroversial: it goes by the name of *Born's rule.* Essentially, one takes the absolute square of the complex number that the wavefunction assigns to each of the configurations to be the likelihood that one will get that configuration on a given run. Every interpretation of quantum theory must vindicate Born's rule as providing at least a very good approximation of the chances for each outcome. But the *practical* utility of Born's rule for making predictions does nothing to solve the ontological problem we have been led into.

That problem is as follows. Irrespective of the use of Born's rule, the end quantum state given by Schrödinger evolution is neither one that simply represents a detector which has fired nor one that represents a detector which has not fired. But, it seems, the actual physical detector, at the end of the experiment, has either fired or not fired. Hence, as John Stewart Bell succinctly put it, 'Either the wavefunction, as given by the Schrödinger equation, is not everything, or it is not right' (Bell 1987: 201).

The idea that the wavefunction is not everything is the claim that the quantum state of a system does not provide a *complete physical description* of the system. Certain important aspects of a physical object may be captured by its wavefunction, but something at least as important is left out: whatever it is that makes a detector which has fired physically different from one which has not. Einstein, Podolsky, and Rosen raised the question 'Can quantum-mechanical description of reality be considered complete?' in an article with that very title, and concluded that it cannot (Einstein *et al.* 1935). Einstein, Podolsky, and Rosen did not themselves offer any explicit account of how the quantum-mechanical description was incomplete, or what it would take to complete it, or how to understand the wavefunction, but it is clear what sort of a job confronts the physicist who denies the completeness of the quantum state. If the wavefunction is not *everything*, then we need to know *what else there is*, and what is the *dynamics* of this extra stuff, and how the wavefunction comes into the story at all. Theories of this kind are commonly called *hidden-variables theories*, the 'hidden' variables being representations of whatever it is the quantum state leaves out of account. The term 'hidden variables' is particularly badly chosen, since these additional variables are supposed to represent whatever physical difference it is which distinguishes a detector which has fired from one which has not, and that difference, far from being hidden, is paradigmatically manifest. But in any case, the hidden-variables interpretations are free to regard the wavefunction, as given by the Schrödinger equation, as right but incomplete.[4]

The two questions which confront a hidden-variables theory, then, are 'What is there beside the wavefunction?' and 'What dynamical laws govern this additional ontology?' Several different such theories are currently on offer, the most famous being that originally suggested by Louis de Broglie and later developed by David Bohm. General algorithms for constructing hidden variables theories have been intensely studied by philosophers, often under the rubric of 'modal' interpretations (van Fraassen 1991; Bub 1998; Dickson 1998; Vermaas 1999). The term 'modal' traces back to the work of Bas van Fraassen, who got some of his ideas from analogies with modal logic, but the philosophical account of modality really plays little role in these theories. (Any incomplete physical description of a system can be considered

[4] A 'hidden variables' theory obviously *could* maintain that the wavefunction, as given by the Schrödinger equation, is *neither* right nor everything: one could both emend the linear evolution and add more variables. In practice, though, this has not been pursued, since the addition of the 'hidden' variables alone (with the right dynamics for them) can do the job.

'modal' in so far as it is consistent with various different completions, and so at best constrains how the system might be, but does not indicate exactly how it is.)

In Bohm's version of quantum *mechanics*, the additional variables are particle locations: particles, in this theory, always have exact locations, and hence form a single unique configuration, even when the wavefunction is spread out all over configuration space. The wavefunction always evolves in accord with Schrödinger's equation. The key to the theory is then a new equation, the *guidance equation*, which specifies how the positions of the particles evolve with time. The guidance equation makes use of the wavefunction: how a particular configuration will evolve is determined, in a simple way, by the form of the wavefunction.[5]

Hidden-variables theories seize the horn of Bell's dilemma which says that the wavefunction is not everything. The other horn is grasped by *collapse* theories. If the wavefunction is a complete representation of the physical world, then, it seems, it must evolve in such a way as to end up representing either a live cat or a dead cat, either a detector which has fired or one which has not. Schrödinger evolution, as we have seen, does not give this result, so the collapse theories must postulate that at least sometimes (if not always) the wavefunction does not evolve in accord with the Schrödinger equation. Of course, the exquisitely precise predictions of the quantum theory have all been derived by using the Schrödinger (or other appropriate linear) equation, and one doesn't want to throw out the baby with the bathwater. So the trick of a collapse theory is to come up with a new dynamics which at least closely mimics Schrödinger evolution for the sorts of small systems to which the quantum formalism can be applied and solved. Since the founders of the quantum orthodoxy, Bohr and Heisenberg, evidently meant to reject the incompleteness of the quantum state, they were committed (at least implicitly) to some such theory.

The most straightforward way to do this, exemplified by the classic account of von Neumann (1955), is to postulate pure Schrödinger evolution most of the time, punctuated by distinctly non-Schrödinger evolution (the 'collapse of the wavefunction') from time to time. The question which faces such a theory is *when* and *how* such collapses occur. The orthodoxy's first line of response is: (1) the collapse occurs when a measurement is made and (2) the state collapses, randomly, with the appropriate probabilities (provided by Born's rule), to a state in which the measured quantity has a definite value (a so-called *eigenstate* of the quantity). This response, though, is just a holding operation until some more rigorous account is provided of exactly what it takes for there to be a measurement, and what determines exactly which quantity is measured. It is here that some of the most extraordinary claims about the implications of quantum theory have been made, under the general rubric 'participation of the observer'. But we should note immediately that nothing in the logic

[5] There are many sources available with extensive discussions of Bohm's theory. Many topics are covered in Bohm and Hiley (1993). Important foundational issues concerning the role of probability in the theory are examined in Dürr *et al.* (1992).

of the physical problem requires a collapse theory to even mention measurements or observers. In the most rigorous formulated collapse theories, of Ghirardi, Rimini, and Weber (1986) and of Perle (1990), the *when* is answered: 'randomly, with fixed probability per unit time' and the *how* roughly 'to a state in which at least one particle has a much more sharply localized position'.[6] I will examine other possible answers to the 'when?' and 'how?' questions when I look at particular issues.

This brief overview of interpretations of the quantum formalism would not be complete without some mention of a notorious attempt to escape between the horns of Bell's dilemma. The idea is to reject both collapse in the dynamics of the wavefunction and the claim that the wavefunction is incomplete: the wavefunction given by pure Schrödinger evolution both is right and is everything. How then to understand the wavefunction of Schrödinger's cat, which has equal portions for the configuration of a live cat and for a dead one, or the wavefunction of our detector, split between regions of configuration space which correspond to the detector having fired and regions which correspond to no detection? The idea is that such a wavefunction represents *both* outcomes having occurred: the cat both survived and died; the detector both fired and did not. Why then does it seem to us that the cat is just plain alive, and that the detector failed to detect anything? Well, that is because the world has somehow split into two non-interacting parts, with different outcomes in each, and we are only aware (owing to the lack of interaction) of one part, the 'world' we now inhabit.

This is the so-called 'many-worlds' interpretation of quantum theory. The interpretation is often associated with the name of Hugh Everett, even though Everett himself never used the term, and, I think, never held the view. Everett called his interpretation the 'relative state' interpretation, and he paid particular attention to a mathematically well-defined object, the state of one part of an entangled system 'relative to' an *arbitrarily specified* state of another part of that system (Everett 1957). The relative state plays no central role in the many-worlds theory. Indeed, Everett's own interpretation is rather obscure on some central points, and so we will leave it out of account. But the many-worlds view deserves some attention.

I don't think that the many-worlds view, as just sketched, can be considered to be a viable interpretation of the quantum formalism-*cum*-practical-rules-of-thumb that we originally set out to understand. For that practical apparatus has, as a central feature, techniques for assigning *probabilities* to events, and it is the astonishing accuracy of those predictions that provide our grounds for taking the theory seriously. So any coherent interpretation which vindicates our use of the theory must hold that there is something those probabilistic predictions are *about*, and that those things have, in experiments, actual *frequencies* which closely

[6] In the continuous localization theories, the *when* for collapse becomes *all the time*, via a stochastic process which, for small systems, closely approximates Schrödinger evolution over normal laboratory timescales.

approximate the probabilities derived from the theory. In the hidden-variables approaches, the probabilities are ultimately about the actual values of the 'hidden' (i.e. manifest!) variables, which variables display some particular distribution over a series of experiments. For the collapse theories, the probabilities are for the collapses to occur in one way rather than another, and again there is, according to the theory, a fact about the frequencies with which various sorts of collapse occur. But on the many-worlds view, it is hard to see what the probabilities are probabilities *for*.

The probabilities are not for the 'hidden' variables to take on certain values, for there are none. They are not for the wavefunction to evolve one way rather than another, since it always evolves in accordance with the Schrödinger equation. According to this theory, what happens when a measurement takes place is that the world splits. But then what could it mean to assign, say, a 20 per cent probability to one outcome over the other, since both will surely occur, and in both there will equally be a descendant of the 'you' who started the experiment. The probabilities are calculated from the amplitude of the wavefunction, *but that amplitude does not play any metaphysical role in the theory.* If the world splits (assuming we can makes sense of that notion), it splits simply because the wavefunction is spread out a certain way in configuration space, and has nothing to do with *how much* of the wavefunction is in the different places.

Some people seem to equate the 'many worlds' talk simply with a no-collapse theory, i.e. they say that there are many worlds simply because the wavefunction is spread out. In this sense, all of the hidden-variables theories are many-worlds theories. But this usage of language is more likely to confuse than enlighten, for in the hidden-variables theory there is only one, unique distribution of values for the additional variables, which distribution corresponds to the world we see. Even when these variables are influenced by 'interfering' parts of the uncollapsed wavefunction, there is but one manifest world involved: the world of the additional variables. These theories have a physically dualist ontology: the physical world contains both the wavefunction and the additional variables, but of the two components it is the *wavefunction* which is hidden, its form and very presence made known only via its effect on the additional variables. So we will leave the many-worlds idea aside, as insufficiently clear to constitute an interpretation of the quantum formalism.[7]

Let's now turn to a series of ontological issues, and see how they come out according to various interpretations of the general form I have been examining.

[7] A somewhat more detailed discussion of the central interpretative problems confronting quantum theory can be found in Maudlin (1995). Remarks on the relation between the quantum formalism and the manifest world are amplified in Maudlin (1997).

1. Issue One: Determinism

Historically, the most widely remarked metaphysical innovation of quantum theory over classical physics is the rejection of determinism in favour of chance. Events such as the decay of a radioactive atom are typically held to be fundamentally random: there is no reason at all that the decay takes place at one time rather than another. Atoms that are physically identical in every respect may nonetheless behave differently. Einstein was resistant to the idea the God plays dice, and his insistence on determinism is taken to be a mark of a reactionary inability to accept the quantum theory.

Things are not quite so simple. Does either the pragmatic quantum formalism or the empirical result of any experiment require us to abandon determinism? No. The pragmatic formalism requires an interpretation, and some of the interpretations posit deterministic laws while others employ fundamentally stochastic dynamics. Further, little can be said in the way of generalization.

The Schrödinger equation itself is deterministic. So any interpretation which does not employ wave collapse at a fundamental level must find its indeterminism elsewhere (if it is to find it at all). As we have seen, theories which forgo collapse already have a problem to deal with: they need some additional physical stuff (beside the wavefunction) if they are to have any hope of modelling the world as we know it. The question of determinism for these theories becomes a question of the dynamics of this additional stuff, the 'hidden' variables.

The possibility of supplying the additional variables with a deterministic dynamics was demonstrated by the de Broglie–Bohm particle mechanics. The guidance equation in that theory is deterministic: given the initial state of the wavefunction of the universe and the initial configuration of the particles (in this case, their initial locations), the laws of the theory allow only one possible history for the universe. And since the theory makes the standard predictions, for example, the two-slit experiment and electron tunnelling and radioactive decay, these sorts of phenomena cannot assure us that the world operates by chance.

It is often thought that the whole *point* of Bohm's theory, and of 'hidden' variables in general, is to restore determinism (indeed, it is common to regard the purpose of the additional variables themselves as providing the hidden cause for an experiment to come out one way rather than another), but this is inaccurate on several counts. First, the main problem to be solved in these theories is not to give an *explanation* of why one result happened rather than another, but rather to have the theoretical resources to describe the experiment as having *had* one result rather than another. That problem is answered in the first place simply by having more than the wavefunction in the physical ontology, irrespective of the dynamics. *Most* hidden-variables theories, including the so-called modal interpretations, postulate *stochastic* dynamics for the additional variables. Bell, who was one of the great advocates of Bohm's theory, suggested a stochastic dynamics for his version of Bohmian field theory (Bell 1987, ch. 19), and Bohm himself was wont to speculate about an indeterministic 'sub-quantum' realm. The goal of Bohm's account was never determinism *per se*; it was clarity and precision in the theory.

What of collapse theories? Most of these do consider the non-Schrödinger evolution of the wavefunction to be fundamentally indeterministic. The original Spontaneous Collapse theory of Ghirardi, Rimini, and Weber (1986) made the collapses out to be discrete, unpredictable events which happen with some fixed probability per particle per unit time. But again, the question of determinism is only tangential to the motives of the enterprise. In moving from the discrete collapses to the 'continuous spontaneous reduction' model of Philip Perle, for example, the sudden reductions have been replaced by a coupling to a background 'white noise' which determines how the reduction occurs (Perle 1990). And what causes the white noise? The theory does not say (and does not need to, for the purposes at hand). It might be generated deterministically as well as stochastically. So we can't say that the quantum theory forces indeterminism on us. Furthermore, the whole issue looks more like a case of spoils to the victor than a fundamental point of contention: if some consideration militates in favour of a specific interpretation, the question of determinism will simply follow suit, and it seems very unlikely that determinism itself will be a decisive consideration. No one would unnecessarily complicate an interpretation either to instigate, or to avoid, deterministic dynamics.[8]

2. Issue Two: Determinateness

Allied to the question of determinism is a slightly different issue, which goes under the rubric 'determinateness of properties'. It is best illustrated by an example which can he raised only from within an interpretation. Suppose one rejects additional variables—the wavefunction is complete—and for simplicity suppose there is only one particle under consideration, so the wavefunction can be pictured without too much imprecision as a field in space. Further suppose, as will typically be the case, that the wavefunction is spread out, with non-zero values over a large region. What are we to say about the *position* of the particle in this situation?

[8] Ironically, mainstream physicists are at least as likely to attribute determinism as indeterminism to quantum theory. The so-called 'information-loss' paradox used in quantum cosmology is founded on the claim that quantum theory does not allow information about the physical state of a system ever to be lost, even if the system is tossed down a black hole. But if there is any fundamental indeterministic collapse of the wavefunction, information is lost all the time, whenever a collapse occurs.

The pragmatic apparatus (i.e. Born's rule) tells us what to predict in this circumstance if we should happen to *look for* the particle by means of, for example, a fluorescent screen. We could only make probabilistic predictions, with the probability for 'finding' the particle in a given location (i.e. the probability that a flash will occur on the screen in a given location) being equal to the squared amplitude of the wavefunction at that location. But what of the particle right before the flash? Did it have any particular position at all?

If the wavefunction is complete, then it obviously is incorrect to say that immediately before the flash the particle was in the vicinity of the part of the screen where the flash later occurred and not elsewhere. If the wavefunction was spread out over a large region, and if all physical facts are determined by the wavefunction, then the most one can say of the particle is that it was spread out. Hence in this sort of interpretation it is misleading to say that one 'found' the particle in a particular location, as if it had been there all along and its true location was merely *revealed* by the screen. According to this sort of theory, the observation does not reveal anything; it is rather an interaction which *creates* a more localized wavefunction from a less localized one. And before the observation, the right thing to say is that the particle had no determinate location at all.

Failure of determinateness in this sense is not the same as failure of determinism. Empedoclean atoms that sometimes randomly swerve in their trajectories are not deterministic, but seem to have, at all times, perfectly determinate properties. (They do not have perfectly determinate *propensities* since their future behaviour is unpredictable, but that failure is laid at the door of the dynamics.) What seems peculiar about the quantum particles is that they *sometimes* have determinate positions (after an observation using a screen, for example) but at other times do not.

But once one gets accustomed to the ontology of this sort of interpretation, this feature no longer seems at all mysterious. If the wavefunction of a particle is complete, then the only sense in which the particle can have a definite location is for the wavefunction to be localized, i.e. to be non-zero only in a relatively small region. That may happen, but even when it does, Schrödinger evolution guarantees that it won't last long. And once the wavefunction has spread out, the most one can say of the particle is that it is in a state which has the propensity, in varying degrees, to cause flashes on screens at various locations.

On the other hand, adopting a different interpretation can completely alter these conclusions. Bohmian particles, for example, always have exact determinate locations, no matter what the wavefunction is, and in Bohmian mechanics flashes on screens are caused by particles which were, immediately before the flash, in the vicinity of that part of the screen. The various 'modal' interpretations differ both on which properties are determinate at a given time and on what makes them determinate. In some interpretations,[9] what properties are determinate changes

[9] I refer to interpretations which use the polar decomposition theorem to pick out a preferred basis, such as Kochen (1985).

through time, depending on the wavefunction; in others it is fixed. If the interpretation is clear and coherent, it will posit a fundamental ontology, which is governed by some dynamical equations. It is then a relatively straightforward matter to analyse a given experimental procedure in terms of this ontology to discover whether the outcome of the experiment is a reliable indication of any pre-existing state of affairs. So once again, one cannot sensibly ask whether a given property of a system is determinate according to quantum theory; one must rather ask about the account of a particular experimental situation given by a particular interpretation of the theory. One has no reason to anticipate, though, that quantum systems will always be in states which assign determinate values to classical properties like position and momentum and energy.

3. Issue Three: The Role of the Observer

Perhaps the most metaphysically intriguing claim associated with quantum theory is the notion that it somehow introduces the observer, and irreducible subjectivity, back into physics at the most fundamental level. While classical physics aspired to a 'God's-eye view' of the universe, a purely objective and 'mechanistic' account of the world, we are sometimes told that quantum physics has rendered any such notion obsolete. At its most radical, this view suggests that it is only by the 'participation' of the conscious observer that the physical universe came into being at all, leaving us with the rather perplexing problem of how conscious observers themselves arose in the first place.[10] Perhaps unsurprisingly, the situation does not looks so dire when viewed through the lenses of our precise interpretations.

One way the observer might be thought to make an entrance into our story is via the so-called 'measurement problem'. A measurement, it is said, requires a measured system and a measuring system, and the measuring system must be some sort of observer, so without observers there are no measurements and so no measurement problem. Conversely, if there is a measurement prohlem, it must arise because of the presence of an observer.

It is true that some of the basic interpretational problems of quantum theory are often presented using measurement operations for illustrative purposes. Above, we

[10] Perhaps the most striking presentation of this view is not any explicit theory, but rather a picture which appears in John Wheeler's 'Law without Law' (Wheeler 1983*a*: 209). The illustration purports to show the universe as a 'self-excited circuit', in which observers who arise long after the big bang somehow impart 'tangible reality' to their own distant past by means of their observations.

considered what happens when a diagonally polarized photon is shot at a vertically oriented polarizer with a photodetector behind it. We saw that if the wavefunction does not collapse, the resulting quantum state will be a superposition of a state in which the detector fires and a state in which it does not, and we asked after the appropriate understanding of such a state. Since the laboratory operation we described constitutes what would normally be called a measurement of the vertical polarization of the photon, one might describe the problem of interpreting the superposition as the problem of understanding a measurement interaction, and hence a measurement problem.

But there are several things that must be immediately noted. The first is that problematic superpositions are not confined to the results of 'measurements'. Schrödinger's cat ends up in a problematic superposition if the wavefunction does not collapse, but the cat does not, in any straightforward sense, measure anything. As Philip Perle has put it, quantum theory does not so much have a measurement problem as a reality problem: we have to figure out how the quantum formalism represents anything at all as happening in the world, not just measurement interactions. Secondly, in order for the problem to arise, one does not need a *conscious* observer. The polarizer-cum-photodetector is not conscious, but would nonetheless normally be taken to constitute a measuring device. Thirdly, and most importantly, since interpretations such as Bohm's and Ghirardi, Rimini, and Weber's collapse theory are able to account for the behaviour of photodetectors and cats (as we take them to behave) without making any mention of consciousness, conscious observers do not need to be introduced in order to make sense of the mathematical formalism.

From the point of view of Bohm and Ghirardi *et al.*, 'measurement' interactions are simply a species of physical interaction like any other, governed by the same basic dynamical laws as everything else. If a system has a particular physical constitution, it may turn out to be a good indicator of something else. Measurement, as a physical matter, requires the existence of a sort of system (the measuring device) so constituted that, after interacting with a target system (the measured system), its state becomes *correlated* with that of the target system. This correlation means that after the interaction the state of the measuring device contains information about the measured system. None of this requires consciousness.

So how did consciousness, and the *human* observer, ever come into the discussion?[11] The only plausible account is that one may be driven to advert to consciousness as an act of desperation. Suppose one wants a collapse theory, and begins to consider under what conditions a collapse occurs. It *cannot* be that just any

[11] Even if consciousness somehow can get into the game, what's to say that it must be human consciousness? Einstein reportedly voiced his scepticism concerning the suppposed effects of conscious observation by saying that he couldn't believe that a mouse could bring about drastic changes in the universe simply by looking at it (this anecdote is recounted by Everett in his thesis; DeWitt and Graham 1973: 116).

interaction causes the wavefunction to collapse: that the wavefunction of an electron does not collapse whenever it interacts with another electron can be verified by experiment, since some observable interference effects depend on the interaction of both parts of the uncollapsed wavefunction. So one may naturally look for some special species of interaction which could cause collapse. In particular, one may naturally come to think that any interaction which can be understood as just electrons and protons and neutrons following the usual physical laws is *not* special, since the basic physics of such interactions does not require collapses to occur, as we have just seen.

At this point the following line of thought takes over. Let's consider, say, an electron which is fed through a Stern–Gerlach device, i.e. a device which 'measures' spin. In particular, consider an electron which is not in an eigenstate of x-spin and is fed into a device which measures x-spin, so that particles with x-spin up exit through the top of the device and particles with x-spin down exit through the bottom. Since our electron is not in an eigenstate, the wavefunction of the particle after the interaction will be a superposition of one in which the electron exits at the top and one in which it exits at the bottom. And we can *experimentally confirm* that at the moment the particle exits, the wavefunction has *not* collapsed into one or the other definite position by recombining the two beams and looking for interference effects (cf. Albert 1992, ch. 1). So interacting with a Stern–Gerlach device does not collapse the wavefunction.

If we do not recombine the beams, then the wavefunction of the electron is a superposition of states with different positions. Now suppose we decide to look for the electron by, say, putting up a fluorescent screen. The basic physics of the interaction of the electron with the screen, producing light, is well understood. There is nothing terribly exotic in this interaction, nothing fundamentally different from, say, the interaction of a single electron with another. So since merely interacting with an electron does not collapse the wavefunction, it is hard to see how interacting with the screen will. The screen should end up in a superposition of have excited electrons in one place and having excited electrons in another. And when the electrons return to their ground states, we should have light in a superposition of having flashed from one place and having flashed in another. And as we trace the career of the light, we understand how it will interact with the eye, and with the photoreceptors in the retina, and none of this require fundamentally new physics. So the retina of someone watching the screen should end up in a superposition of having one set of rods fire and having another set of rods fire.

Beyond this point, it is not accurate to say that we have a clear understanding of how things work. But we think that the passage of the neural signal down the optic nerve is a matter of simple chemistry, as is the way that the firing of one set of neurons causes the firing of another set. We think that all of this brain activity can more or less be understood in terms of chemistry, and that chemistry can more or less be understood in term of physics, and that although this gets very complicated it does not involve any fundamentally new physics, and so nothing which could do anything as dramatic as set off collapse of the wavefunction.

If one adheres to this line of thought, then the collapse can only be caused by something fundamentally new, by something that we don't have any idea, even more or less, how to understand in terms of the laws of physics and chemistry. And of course, there *is* something mysterious which fits this definition: the general relation between the physical state of the brain and conscious experience. Long before quantum theory, the mind–body problem was recognized as such an explanatory gap. So if we are looking for a trigger for wave collapse with the idea that the trigger must be some fundamentally new sort of interaction, and if we have already concluded that the interaction of mind and body is fundamentally unlike the interaction of body and body, then it is natural to locate the mysterious collapse process here. If we already have something we don't understand, then it seems economical to assimilate it to something else we don't understand. It is not that something about wave collapse could obviously explain consciousness, or that something about consciousness could obviously explain wave collapse, but rather that the explanatory gap between body and mind yawns wide enough to engulf the problem of collapse without so much as a tremble.

If this is what has led theorists to link quantum mechanics to consciousness, then two observations are in order. The first is that the whole line of argument is based on an unwarranted supposition. That supposition is that there must be some special kind of circumstance or interaction (a 'measurement' or 'observation') which 'triggers' collapse of the wavefunction. As the example of Ghirardi, Rimini, and Weber's theory shows, no such trigger is needed: in that theory, collapses happen at random, with a fixed probability, and are not particularly associated with any kind of interaction. The second is that even if one demands some particular circumstance for collapses, there is nothing in the *physics* which points to consciousness or mind as the key. All we know experimentally is that certain sorts of interactions do not collapse the wavefunction, but the differences between those interactions and typical 'measurement' interactions are manifold. Roger Penrose, for example, has speculated that collapses are tied not to consciousness but to *gravitation*: collapses occur when the superposed states differ enough in their gravitational structure (Penrose 1994: 339 ff.). In a way, this appeals to another explanatory gap: just as we don't have a good theory of the mind–body interaction, we don't have a good quantum theory of gravity. The point is that even if we accept that some special circumstance plays a role in collapse, there is absolutely nothing in the *phenomena* which points to consciousness rather than, say, gravity as the special ingredient. So unless one is already inclined to put the observer as the centre of one's theory, there is nothing in quantum physics to suggest that one do so.

Appeals to consciousness are not restricted to collapse theories. Non-collapse theories, as we have seen, must appeal to some ontology beyond the wavefunction in order to solve the measurement problem. The hallmark of these 'additional variables' is that they have determinate values even when the quantum state is not in an eigenstate of the corresponding operator. Probabilities appear in such theories as the likelihood, given the quantum state, that these additional variables take some particular value. In Bohmian mechanics, and Bohmian field theory, and the various 'modal' interpretations, the additional variables are purely physical: they have no intrinsic connection to consciousness. But in the Many-Minds theory of David Albert and Barry Loewer (1988), the additional variables, which always have determinate values, are conscious states. Like the variables in the modal interpretations[12] and in Bell's version of Bohmian field theory, the conscious states evolve indeterministically, giving a straightforward way to understand the probabilities of the theory.

It may come as a surprise that probabilities are easy to understand on the Many-Minds view given the argument above that the probabilities cannot be interpreted in the Many-Worlds view, rendering it unacceptable. It is important to note that the *source* of the multiplicity in each case is entirely different. The many 'worlds' are generated, as it were, by fission of a single parent world, so that no meaning can be ascribed to attaching different probabilities to the offspring, all of which are certain to be produced. In the Many-Minds theory, minds never fission: each mind has a single determinate history, such that we can ascribe definite frequencies to the apparent results of experiments (as perceived by the mind). Indeed, as far as the interpretation of probability goes, the multiplicity of minds plays no role whatsoever: the stochastic dynamics governs each single mind individually, and the so-called Single-Mind theory suffices for this purpose (Albert and Loewer 1988: 205). The reason to associate many minds rather than just one with each body concerns a desire to maintain something akin to the supervenience of the mental on the physical (although the conscious state of any particular mind does not supervene on the physical state of the associated body, and had better not, lest the wavefunction again become complete!), and with the desire to ensure that one is not typically misled into thinking that there are other minds with certain contents when there are none. Neither of these motivations come from physics *per se*, and so the multiplicity of minds plays no role in the interpretation of the purely physical aspects of the theory.

We can therefore say of the Many-Minds theory as we did of the collapse-inducing-consciousness theory that there are no *physical* considerations which militate in favour of appeals to consciousness here. The belief that consciousness should play a central role in an interpretation of quantum theory must ultimately rest on views about consciousness which are imported into the physics rather than being derived from it.

[12] This remark does not apply to van Fraassen's (1991) interpretation, where the dynamics for the additional variables (the 'value state') is not explicit.

4. Issue Four: Uncertainty and Complementarity

In light of the results of the first three issues, we can deal with this one expeditiously. All interpretations accept use of the wavefunction and Born's rule as a good device for making predictions about systems. These predictions are typically probabilistic, but for each observable characteristic, such as position or momentum, there are special states (the eigenstates for the observable) which allow one to predict with certainty what the outcome of an observation will be. It is a simple mathematical fact that the eigenstates, for example, the position of a particle in some dimension are not eigenstates for its momentum in that dimension, and vice versa. So there exist no quantum states which allow for certain prediction of both the outcome of a position measurement and the outcome of a momentum measurement. Furthermore, there is a quantifiable relation between the uncertainties associated with the measurements: the more certain one is about how one measurement will come out, the less certain one must be about the other. This is the Heisenberg Uncertainty Principle, and the relevant pairs of observables are called complementary.

The question which naturally arises is whether the Uncertainty Principle states a limitation on our *knowledge* or a more fundamental limitation on *the world itself*. For convenience, let's grant for the moment that all there is to having a determinate position or determinate momentum is to be disposed to produce a certain outcome in the appropriate experimental situation (a position or momentum measurement). Then the question is: is it possible for a particle to be disposed to give a particular result for both a position and momentum measurement, and so to have determinate values for both, but we just can't know of both dispositions simultaneously, or is it more fundamentally that no particle can have both dispositions at the same time?

Unsurprisingly, the answer depends on the interpretation one adopts. If the wavefunction is complete, then the more fundamental, ontological condition applies: no particle can have both dispositions at once. This is a consequence of the stochastic dynamics of the system, the fact that unless the system is in the appropriate eigenstate, nothing at all determines what the outcome of the experiment will be. In this case, the relevant uncertainty is uncertainty about how experiments will come out, but not uncertainty about the present state of the system: in knowing the wavefunction, one knows all there is about the system, and is not *ignorant* of anything. The Uncertainty Principle is then a limitation on one's knowledge of the present state of a system only in a Pickwickian sense: the limitation is in the facts to be known, not in our knowledge of any facts.

On the other hand, in a deterministic theory such as Bohm's, the Uncertainty Principle must be epistemic rather than ontic. Since the theory is deterministic, the outcome of any precisely specified experiment must be determined by the initial state of the system and apparatus, so if one knew enough about the system and apparatus one could predict with certainty how a particular position measurement or momentum measurement or any measurement would come out.[13] Conversely, our inability to predict the exact outcome of any experiment (as codified in Born's rule) must be a consequence of our ignorance of certain physically relevant facts about the system.

The natural question which then arises is why this sort of ignorance would be enforced on us, why we could not find out the relevant facts which would allow us to make predictions more precise than Born's rule allows. The surprisingly satisfying answer is that as physical objects ourselves, our ability to gather information about the world is constrained by physical laws. The very deterministic dynamics which ensures that there is a fact about how a particular particle would react to any possible experiment also precludes our coming to those facts by interacting with the system. Proving this requires a careful consideration of how one physical system can gather information about another, i.e. how the state of one can become correlated with the state of the other by means of a physical interaction (see Dürr *et al.* 1992). But it is exactly by taking observers seriously as *physical objects, subject to the laws of physics* (rather than as something outside or distinct from 'purely mechanical' physical systems), that Bohm's theory explains the Uncertainty Relations.

5. Issue Five: Quantum Logic

Perhaps the most intriguing claim about quantum theory is that it provides empirical grounds for revising *logic itself.* The general idea that logic could be become an empirical matter was advanced by Quine in 'Two Dogmas of Empiricism' (Quine 1951), and the more specific proposal to interpret the meet and join operation on the lattice of quantum propositions as the 'true' meaning of 'and' and 'or' traces back in the physics literature to Birkhoff and von Neumann (1936) and has been considered in the philosophical literature in Putnam (1969) among many others. The proposal is given precise meaning by associating quantum propositions with subspaces of Hilbert space, and understanding the 'conjunction' of two propositions as the intersection of the two spaces and the 'disjunction' of two propositions as their span. The physical state of the world is represented by a vector in the Hilbert space,

[13] One has to be a bit cautious here. The outcome of any precisely described experiment could be predicted; that does not mean that one could assign a value to every mathematically defined 'operator': in Bohm's theory, physically different experimental set-ups which would be associated with the same mathematical operator could evoke different outcomes. See Albert (1992: 153).

and a proposition is true just in case the vector lies in the subspace associated with the proposition. It immediately follows that a 'disjunction' can be true even though neither 'disjunct' is true, since a vector can lie in the span of two subspaces without lying in either subspace.

Even more striking is the failure of distributivity of the lattice of quantum propositions. Take any three non-collinear vectors **A**, **B**, and **C** such that **C** lies in the subspace spanned by **A** and **B**. Let '$\vee$' represent the join of two subspaces and '$\wedge$' represent the meet. Then $(\mathbf{A} \vee \mathbf{B}) \wedge \mathbf{C} = \mathbf{C}$, while $(\mathbf{A} \wedge \mathbf{C}) \vee (\mathbf{B} \wedge \mathbf{C}) = \mathbf{0} \vee \mathbf{0} = \mathbf{0}$. So if the vector which represents the state of the system is $\mathbf{C}$, $(\mathbf{A} \vee \mathbf{B}) \wedge \mathbf{C}$ is true while $(\mathbf{A} \wedge \mathbf{C}) \vee (\mathbf{B} \wedge \mathbf{C})$ is not only false but necessarily false. Distributivity of meet over join fails. If we interpret meet and join as 'and' and 'or', then in this 'logic' de Morgan's law can fail. We are therefore outside the domain of classical logic.

A more intuitive presentation of the basic idea can be given by considering the standard two-slit experiment. Surely, it seems, any particle which gets to the screen must have passed through one slit or the other. But if it passed through the top slit, it might show up anywhere on the screen (a beam of particles shot through the top slit will not form interference bands), and if it passed through the bottom slit, it might show up anywhere on the screen (similar reasoning), but in fact there are places on the screen (the dark bands) where it will not show up. So it's not true that it went through the top slit and not true that it went through the bottom slit even though it is true that it went through the top or the bottom. Hence a disjunction can be true even though neither disjunct is true, so classical logic fails.

If a proposition is true just in case the quantum state is an eigenstate of the appropriate operator, then we can understand this result as follows: the wavefunction of the particle is not in an eigenstate of being located at the top slit (since not all of it passes through the top slit), nor is it in an eigenstate of being located at the bottom slit (for similar reasons), but it is in an eigenstate of being located at the union of the two slits. In this sense it passes through 'the top or the bottom' without passing through the top or passing through the bottom. So again, classical logic fails.

Or rather, it is now perfectly clear that classical logic does *not* fail. One would only think it did if one made the egregious mistake of thinking that the proposition that the particle passed through the *union* of two regions is the same as the disjunction of the proposition that it passed through the first with the proposition that it passed through the second. But the two are not equivalent. It may be true that the Rocky Mountains are located entirely in the union of the United States and Canada, but that is not the disjunction of the proposition that they are located entirely in the United States with the proposition that they are located entirely in Canada. Facts about unions are not just disjunctions of corresponding facts about the parts which comprise them.

In some cases, of course, facts about unions do correspond to disjunctions of facts about their parts. If a *pointlike* object is located in the union of the United States and Canada, then the object is either located in the United States or it is located in Canada. So if electrons were pointlike particles, always having a determinate position, then an electron could only pass though the union of the slits by passing through one slit or the other. But, as we have seen, if the wavefunction is complete, then the electron does not have a single determinate location, so one is not entitled to the inference.

Similarly, the right conclusion to draw from the non-distributivity of the lattice of quantum propositions is simple: the 'meet' and 'join' of two propositions on the lattice are not the conjunction and disjunction of those propositions. There may be circumstances in which the meet has the same truth-value as the conjunction, and the join the same truth-value as the disjunction, and indeed this might typically be so in circumstances where quantum effects are absent, but the very examples which are supposed to convince us that classical logic fails really only demonstrate that quantum 'logic' isn't *logic*, i.e. isn't an account of conjunction and disjunction. The two-slit experiment casts no more doubt on classical logic than do the Rocky Mountains.

What of theories in which the wavefunction is not complete? What, for example, of Bohm's theory, in which the electron *is* pointlike and always *does* have a single determinate location? In that theory it is true that each electron passes through the union of the two slits, but equally true that each electron passes through either the top slit or the bottom slit. Indeed, according to Bohm's theory one can tell which slit the particle passed through even though the collection forms interference bands on the screen: all the particles on the upper half of the screen passed through the upper slit and all of those on the lower half through the lower.[14]

But if each particle goes through one particular slit, why do interference bands form? After all, if we block off either slit, then the interference bands disappear, so *something* must be sensitive to the fact that both slits are open. If the particle only goes through one slit, how can it 'know' that both are open?

The answer, of course, is that the form of the *wavefunction* is different when only one of the slits is open rather than both. The wavefunction interacts with both slits even though the particle itself only passes through one. This requires thinking of the wavefunction as something different from the particle itself, but does not require any alteration in logic.

[14] Diagrams of the trajectories of Bohmian particles in this case have been produced; see, e.g. Bohm and Hiley (1993: 33).

6. Issue Six: Entanglement and Non-Locality

So far, the metaphysical results I have surveyed may seem disappointingly modest. There are essentially three points at which the ontology of an interpretation of quantum theory may depart from that of classical physics. One is in the acceptance of probabilistic dynamics at a fundamental level. This introduces objective chance into the theory, and therefore requires that one take dispositions seriously. The consequences of this are relatively straightforward, though interestingly constrained by the mathematical structure of the theory, as revealed in the Uncertainty Principle. We have also seen that this indeterminism is not forced on the theory directly by empirical results, since deterministic interpretations such as Bohm's can handle the archetypal quantum phenomena. A second point of ontological innovation may be introduced if one ties the collapse of the wavefunction or the nature of the additional variables to, for example, consciousness. But this sort of move is completely speculative, with no foundation in experiment at all. The third source of ontological innovation lies in the fact that the wavefunction of a system is a vector space defined on configuration space rather than on physical space. I will now turn to the implications of this circumstance.

The peculiar characteristics of the wavefunction are most easily illustrated by means of a particular state of two spin-$\frac{1}{2}$ particles (such as electrons), the so-called *singlet* state. The spin of a spin-$\frac{1}{2}$ particle can be measured by means of an inhomogeneous magnetic field oriented in a given direction (a Stern–Gerlach device): when passed through the field, the particle will be deflected in either one direction ('up') or the other ('down'). If we orient the field in the x-direction, then we measure x-spin, if in the y-direction, y-spin, and so on. A single particle can be in an eigenstate for spin in any direction, i.e. can be in a state in which it is disposed with certainty to be deflected in a certain way by the field. So among the quantum spin states available to a particle is a state in which it is certain to go up if the x-spin is measured, a state we represent by '$|x\uparrow\rangle$'. Similarly, a particle can be in the state $|x\downarrow\rangle$, in which an x-spin measurement will certainly deflect it down. It can be in the state $|z\uparrow\rangle$, in which it is certain to go up if spin in the z-direction is measured, and so on. It turns out to be a mathematical fact (in the usual representation) that $|z\rangle = \frac{1}{\sqrt{2}}|x\uparrow\rangle + \frac{1}{\sqrt{2}}|x\downarrow\rangle$, so a state with definite z-spin cannot have definite x-spin. In fact, if the z-spin can be predicted with certainty, the results of an x-spin measurement will be completely random. This is an example of an uncertainty relation.

If we have a pair of particles, then among the quantum states available are states where the first particle and the second particle each have given single-particle spin states, such as $|z\uparrow\rangle_1|z\uparrow\rangle_2$, $|z\uparrow\rangle_1|z\downarrow\rangle_2$, and $|z\uparrow\rangle_1|x\uparrow\rangle_2$ (in obvious notation). This is no surprise. But since the space of quantum states is a vector space, we can also form *superpositions* of states like these, by weighting different states by a (complex) coefficient and adding them. The particular state we will be concerned with, the singlet state, is a superposition of the state $|x\uparrow\rangle_1|x\downarrow\rangle_2$ and the state $|x\downarrow\rangle_1|x\uparrow\rangle_2$, namely:

$$\frac{1}{\sqrt{2}}|x\uparrow\rangle_1|x\downarrow\rangle_2 - \frac{1}{\sqrt{2}}|x\downarrow\rangle_1|x\uparrow\rangle_2.$$

What is this state like?

The singlet state is a superposition of two states in each of which the particles have opposite x-spins: in one, particle 1 has x-spin up and particle 2 x-spin down, and vice versa in the other. It is not surprising, then, that the quantum formalism makes a prediction with certainty: if x-spin is measured on both particles, one of the particles will have x-spin up and the other x-spin down. It is also not surprising that the quantum formalism does not predict with certainty *which* particle will be up and which down. Indeed, the formalism (Born's rule) ascribes a 50 per cent probability for each outcome, and hence a 50 per cent probability for either of the particles to display x-spin up if measured. Neither of these facts is surprising by itself, but taken together they are quite puzzling.

Suppose that the wavefunction is a complete physical description of the pair of particles. Then the pair can have opposite x-spins (i.e. be disposed with certainty to give opposite x-spin results if measured), even though neither particle has a determinate x-spin (i.e. a sure-fire disposition to react one way or another to an x-spin measurement). It sounds incoherent to say that neither particle has an x-spin but nonetheless their (non-existent?) x-spins are correlated, but the claim can be made clear sense of in terms of dispositions to respond to various measurements.

It is a mathematical fact that the singlet state

$$\frac{1}{\sqrt{2}}|x\uparrow\rangle_1|x\downarrow\rangle_2 - \frac{1}{\sqrt{2}}|x\downarrow\rangle_1|x\uparrow\rangle_2$$

can equally well be written in terms of z-spin as

$$\frac{1}{\sqrt{2}}|z\uparrow\rangle_1|z\downarrow\rangle_2 - \frac{1}{\sqrt{2}}|z\downarrow\rangle_1|z\uparrow\rangle_2.$$

So what was said for x-spin holds *mutatis mutandis* for z-spin: each particle has a 50 per cent chance of going up or going down if z-spin is measured, but if both z-spins are measured, the results for the two particles are sure to be opposite.

There are several curious features of this sort of *entangled* state of the two particles. From the perspective of fundamental metaphysics, the most important point is that the state seems to exhibit an irreducible form of *holism*. For if we consider, say, particle 1 on its own, and characterize how it is disposed to respond to spin measurements, we can say that is has a 50 per cent chance of going up or down if x-spin is measured, or if z-spin is measured, or indeed if spin in any direction is measured, and the same for particle 2. But knowing all there is to know about each particle individually (in this sense) does not suffice to tell us all there is to know about the pair. For it is a fact about the pair that they are disposed to give opposite results if spins in the same direction are measured, but this fact does not follow from the totality of dispositions of each particle taken separately. Formally, the quantum formalism would ascribe a particular *mixed state* to each particle, but knowing that each particle is in this mixed state is not enough to determine that the pair is in the singlet state.

The failure of the quantum state of the whole to supervene on the quantum states of the parts is most strikingly illustrated by the so-called $m = 0$ triplet state:

$$\frac{1}{\sqrt{2}}|x\uparrow\rangle_1|x\uparrow\rangle_2 - \frac{1}{\sqrt{2}}|x\downarrow\rangle_1|x\downarrow\rangle_2.$$

If one measures the x-spin of either particle in this state, again the quantum formalism ascribes a 50 per cent chance to each possible outcome, and similarly for z-spin, and spin in every other direction. That is, the mixed state ascribed to each particle in the $m = 0$ triplet state is *identical* to the mixed states ascribed to the particles in the singlet state. Nonetheless, the singlet differs from the $m = 0$ triplet. The difference, however, can only be revealed by a *global measurement* made on *both* particles, and not by any possible *local* measurement made on one particle. For if we measure the x-spins of both particles in the $m = 0$ triplet state, they are certain to give the *same* result rather than opposite results, although, of course, half the time they will both be spin up and half the time spin down. So the individual particles in the $m = 0$ triplet state are indistinguishable by any measurement from their counterparts in the singlet state: in so far as the individual particles have quantum states at all, they are identical. But the wholes of which they are parts are nonetheless in physically distinct states, as can be verified by a single global measurement. The quantum state of a whole therefore does not supervene on the states of its parts, exhibiting a form of holism.[15]

Here is another peculiarity about these entangled states. We have said that in the singlet state each particle has a 50 per cent disposition to go, for example, up if x-spin is measured. But suppose we measure the x-spin of particle 1 and it happens to go up. Since we are certain that x-spin measurements on *both* particles are certain to give opposite results, once we have measured particle 1 we can be certain that if the x-spin of particle 2 is measured, particle 2 will go down. That is, after the measurement performed on particle 1, the physical dispositions of particle 2 have changed. And, indeed, the measurement performed on particle 1 will change the *quantum state* assigned to particle 2 via the collapse of the wavefunction. Furthermore, this holds no matter how far apart the two particles are: measuring one will collapse the

[15] A more detailed discussion of this point appears in Maudlin (1998).

wavefunction, and the collapse will change the quantum state ascribed to the other particle. This is what Einstein memorably called *spooky action-at-a-distance*.

All of our discussion so far has been conducted under the assumption that the wavefunction is a complete physical description of the pair of particles. We have already seen that any such theory must employ wave collapse to solve the measurement problem. What we are now seeing is that since the wavefunction is defined on the configuration space of the system, and since every configuration includes the states of all the parts of the system, the collapse can change the quantum states of all the parts, no matter where they are located. So if collapses can be triggered by interaction with single parts, those interactions can have global effects on all the parts, even those which seem to be distant and unconnected.

At this point the reader ought to feel that some very strong, and odd, metaphysical conclusions are being drawn from empirical results which seem fairly unremarkable. After all, all we have said is that if we prepare a bunch of pairs of particles in the singlet state, and then measure their x-spins, or their z-spins, or spins in any other direction, we will find the following two results:

(1) In the long term, the spin measurements in any direction will yield 'up' results half the time and 'down' results half the time.
(2) Whenever we measure spins in the same direction on a given pair, the results will be opposite.

It might well occur to one that this sort of empirical result can be easily obtained without any sort of metaphysical innovation at all. Suppose, for example, that there are really two sorts of pairs of particle produced when we perform the physical operation we call 'making a pair in the singlet state': sometimes we produce pairs in which particle 1 has x-spin up and particle 2 x-spin down (i.e. each particle always has a sure-fire disposition to react in a certain way to an x-spin measurement), and sometimes we produce pairs in which particle 1 is disposed to go down and particle 2 up.[16] If we happen to produce the first sort of pair about half the time and the second sort about half the time, then the empirical results will be as we have described, but no metaphysical funny business is needed. Each individual particle has a sure-fire disposition all along, and there is no 'spooky action-at-a-distance'. On this picture, 'collapse of the wavefunction' is a purely epistemic, rather than ontic, affair: when we measure the first particle, we simply *find out* whether we have created a pair of the first type or a pair of the second type. Our *knowledge* of the distant particle changes when we make a local measurement, but the particle itself is physically unchanged.

[16] I am describing these dispositions verbally rather than by using quantum states since we are not supposing that the true physical state of the particle can be captured by any quantum state. We don't want to assume, for example, that the true states are subject to the uncertainty relations.

This sort of explanation is so simple and prosaic that it seems at first glance perverse to stick to the idea that the wavefunction is complete. For all that we need to do to accept the prosaic explanation is to accept that the quantum state of the pair is not a complete physical description: quantum theory ascribes the same state to pairs of particles that are physically different, namely pairs where particle 1 is disposed to show x-spin up and particle 2 x-spin down, and pairs where it is the opposite. Einstein himself, of course, thought that the spooky action-at-a-distance was physically unacceptable, so that all of this talk about the collapse of the wavefunction just showed that the wavefunction could not be complete. As I mentioned above, the famous paper of Einstein, Podolsky, and Rosen is entitled 'Can quantum-mechanical description of reality be considered complete?', and their answer is, simply, no. From Einstein's point of view, the insistence on the completeness of the quantum state was perverse and unjustified in light of the attendant commitment to action-at-a-distance and the availability of alternative explanations that preserve contact action. And so things remained for thirty years.

The breakthrough in the debate occurred when John Stewart Bell proved his famous theorem. For although Einstein was right about the particular sorts of global measurements he considered—like the measurements considered above, they admit of a simple local explanation—Bell realized that there are other sorts of global measurements one can make. One can, for example, measure the x-spin of one particle and the z-spin of the other, or, more generally, one can measure the spins in arbitrary directions on each side. And what Bell showed is that if one considers the totality of these sorts of measurements, *no local theory can replicate the predictions of quantum mechanics*. That is, even if one denies the completeness of the wavefunction, one cannot get rid of the spooky action-at-a-distance and recover the quantum predictions.[17]

In Bohm's theory, for example, the wavefunction is not complete, and particles always have determinate locations, but still a measurement carried out on one particle can influence other entangled particles, no matter how distant. This influence is mediated by the wavefunction which, as we have seen, is irreducibly holistic. One must always bear in mind that even though Bohm's theory does not take the wavefunction to be complete, it does take it to represent a serious, irreducible part of reality. And even though the wavefunction in Bohm's theory does not itself collapse, its dynamical role allows it to underwrite action at a distance.

So the deepest metaphysical innovations of the quantum theory lie not in indeterminism, or in complementarity, or in the uncertainty relations, or in the role of observation, or in emendations to logic. The deepest metaphysical innovation lies in the holistic nature of the wavefunction, and the fact that the quantum state

[17] There are many accurate non-technical presentations of Bell's theorem in the literature, e.g. d'Espagnat (1979); Mermin (1981); Herbert (1985, ch. 12); Maudlin (1994, ch. 2).

of an entangled system cannot be recovered from the quantum states of its parts. Furthermore, the holism of the wavefunction, together with the dynamics that govern it (if one accepts collapses) or the dynamics it governs (in a hidden-variables theory), imply the existence of action-at-a-distance. And this action is not merely a theoretical posit: it has direct empirical consequences, namely violations of Bell's inequality, which cannot be predicted without it.

One might well wonder (as Einstein would have) whether this action-at-a-distance could be reconciled with Relativity, since the action has to act faster than light. This is an interesting question, and might even make a topic for a book (cf. Maudlin 1994).

References

Albert, D. (1992). *Quantum Mechanics and Experience.* Cambridge, Mass.: Harvard University Press.

——and B. Loewer (1988). 'Interpreting the Many-Worlds Interpretation'. *Synthese*, 77: 195–213.

Bell, J. S. (1987). *Speakable and Unspeakable in Quantum Mechanics.* Cambridge: Cambridge University Press.

——(1990). 'Against "Measurement"', in A. I. Miller (ed.), *Sixty-Two Years of Uncertainty.* New York: Plenum Press, 17–31.

Birkhoff, G., and J. von Neumann (1936). 'The Logic of Quantum Mechanics'. *Annalen der Mathematik*, 37: 823.

Bohm, D., and B. J. Hiley (1993). *The Undivided Universe.* London: Routledge.

Bub, J. (1997). *Interpreting the Quantum World.* Cambridge: Cambridge University Press.

Cushing, J., and E. McMullin (eds.) (1989). *Philosophical Consequences of Quantum Theory.* Notre Dame, Ind.: Notre Dame University Press.

d'Espagnat, B. (1979). 'Quantum Theory and Reality'. *Scientific American*, 241: 158–70.

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

Dickson, M. (1998). *Quantum Chance and Non-Locality.* Cambridge: Cambridge University Press.

Dürr, D., S. Goldstein, and N. Zhangi (1992). 'Quantum Equilibrium and the Origin of Absolute Uncertainty'. *Journal of Statistical Physics*, 67: 843–907.

Einstein, A., B. Podolsky, and N. Rosen (1935). 'Can Quantum-Mechanical Description of Reality be Considered Complete?' *Physical Review*, 47: 777–80. Repr. in Wheeler and Zurek (1981*b*).

Everett, H., III (1957). 'Relative State Formulation of Quantum Mechanics'. *Review of Modern Physics*, 29: 454–62. Repr. in Wheeler and Zurek (1981*b*).

Ghirardi, G. C., A. Rimini, and T. Weber (1986). 'Unified Dynamics for Microscopic and Macroscopic Physics'. *Physical Review*, D34: 470–91.

Healey, R., *The Philosophy of Quantum Mechanics: An Interactive Interpretation.* Cambridge: Cambridge University Press.

Herbert, N. (1985). *Quantum Reality.* New York: Anchor Press, Doubleday.

Hughes, R. (1989). *The Structure and Interpretation of Quantum Mechanics.* Cambridge, Mass.: Harvard University Press.

Kochen, S. (1985). 'A New Interpretation of Quantum Mechanics', in P. Lahti and P. Mittelstaedt (eds.), *Symposium on the Foundations of Modern Physics.* Singapore: World Scientific, 151–70.

Lockwood, M. (1989). *Mind, Brain and the Quantum.* Oxford: Oxford University Press.

Maudlin, T. (1994). *Quantum Non-Locality and Relativity.* Oxford: Basil Blackwell.

—— (1995). 'Three Measurement Problems'. *Topoi*, 14: 7–15 .

—— (1997). 'Descrying the World in the Wavefunction'. *The Monist*, 80: 2–23.

—— (1998). 'Part and Whole in Quantum Mechanics', in E. Castellani (ed.), *Interpreting Bodies.* Princeton: Princeton University Press, 46–60.

Mermin, D. (1981). 'Quantum Mysteries for Everyone'. *Journal of Philosophy*, 78: 397–408.

Miller, A. (ed.) (1990). *Sixty-Two Years of Uncertainty.* New York: Plenum Press.

Penrose, R. (1994). *Shadows of the Mind.* Oxford: Oxford University Press.

Perle, P. (1990). 'Toward a Relativistic Theory of Statevector Reduction', in A. I. Miller (ed.), *Sixty-Two Years of Uncertainty.* New York: Plenum Press, 193–214.

Putnam, H. (1969). 'Is Logic Empirical?' *Boston Studies in the Philosophy of Science*, 5: 199–215. Repr. as 'The Logic of Quantum Mechanics', in Putnam (1975: 174–97).

—— (1975). *Mathematics, Matter and Method.* Cambridge: Cambridge University Press.

Quine, W. (1951). 'Two Dogmas of Empiricism'. *Philosophical Review*, 60: 20–43.

Redhead, M. (1987). *Incompleteness, Nonlocality, and Realism.* Oxford: Clarendon Press.

van Fraassen, B. (1991). *Quantum Mechanics: An Empiricist View.* Oxford: Oxford University Press.

Vermaas, P. (1999). *A Philosopher's Understanding of Quantum Mechanics: Possibilities and Impossibilities of a Modal Interpretation.* Cambridge: Cambridge University Press.

von Neumann, J. (1955). *Mathematical Foundations of Quantum Mechanics*, trans. R. T. Beyer. Princeton: Princeton University Press.

Wheeler, J. (1981*a*). 'Law without Law', in Wheeler and Zurek, 182–213.

—— and W. Zurek (eds.) (1981*b*). *Quantum Theory and Measurement.* Princeton: Princeton University Press.