JEFFREY A BARRETT

THE SUGGESTIVE PROPERTIES OF QUANTUM MECHANICS WITHOUT THE COLLAPSE POSTULATE

ABSTRACT. Everett proposed resolving the quantum measurement problem by dropping the nonlinear collapse dynamics from quantum mechanics and taking what is left as a complete physical theory. If one takes such a proposal seriously, then the question becomes how much of the predictive and explanatory power of the standard theory can one recover without the collapse postulate and without adding anything else. Quantum mechanics without the collapse postulate has several suggestive properties, which we will consider in some detail. While these properties are not enough to make it acceptable given the usual standards for a satisfactory physical theory, one might want to exploit these properties to cook up a satisfactory no-collapse formulation of quantum mechanics. In considering how this might work, we will see why any no-collapse theory must generally fail to satisfy at least one of two plausible-sounding conditions.

1

The standard theory of quantum mechanics tells us that there are two ways the state of a physical system might evolve.[1] If no measurement is made of a system S, its time-evolution is described by the continuous linear dynamics. If S's state at time t_0 is given by $|\psi(t_0)\rangle_S$, then its state at time t_1 will be given by $\hat{U}(t_0, t_1)|\psi(t_0)\rangle_S$, where $\hat{U}(t_0, t_1)$ is a unitary operator that depends on the energy properties of S. But if a measurement is made, then S instantaneously and nonlinearly jumps into an eigenstate of the observable being measured, a state where S has a determinate value for the property being measured. If the initial state is given by $|\psi\rangle_S$ and if $|\chi\rangle_S$ is an eigenstate of the observable O, then the probability of S collapsing to $|\chi\rangle_S$ is equal to $|\langle\psi|\chi\rangle|^2$, the magnitude of the projection of S's premeasurement state onto the eigenstate.

The linear and nonlinear dynamics are mutually incompatible in that they cannot both correctly describe the time-evolution of the same system at the same time. The standard theory, however, fails to tell us what constitutes a measurement, so it does not provide a clear criterion for when to apply the linear dynamics and when to apply the nonlinear dynamics. Further, there are at least in principle empirical consequences to any criterion one might specify for when to apply one or the other of these dynamics. And finally, any criterion one might specify

Erkenntnis 41: 233–252, 1994.
© 1994 *Kluwer Academic Publishers. Printed in the Netherlands*

for what constitutes a measuring device or other nonlinear system is bound to look *ad hoc* unless one can also give some explanation of how systems constructed entirely from fundamental systems that each apparently follow a perfectly linear dynamics end up following a nonlinear dynamics. In short, the standard theory of quantum mechanics is either logically inconsistent or incomplete in an empirically significant way. This is the quantum measurement problem – a resolution would require one to provide a satisfactory alternative to the standard theory.

A resolution was proposed by Hugh Everett in 1957. Everett's idea was to suppose that measurements, just like every other sort of interaction, are perfectly linear. He thus suggested simply dropping the nonlinear collapse dynamics from the standard theory and taking the resulting theory as complete. An immediate reaction might be to argue that the collapse dynamics is necessary for the empirical adequacy of quantum mechanics. Without a collapse, measuring devices would in general report a superposition of results, and presumably no one has ever seen a measuring device, or any other ordinary macroscopic object, in a superposition. But this may be too fast – after all, it is not entirely clear what a superposition would look like.

Given the assumption that all interactions are perfectly linear, Everett believed that he was able to provide some insight into what it would be like to observe a system that is not in an eigenstate of the observable being measured. He argued that an observer might measure a system in a superposition of eigenstates of the observable being measured, end up in a superposition of recording the results corresponding to these various eigenstates, but nonetheless have the *subjective experience* of an ordinary, determinate result to the measurement. While it is not entirely clear how Everett's relative-state formulation of quantum mechanics was supposed to work, he was convinced that quantum mechanics without the collapse postulate provided a "more general and complete formulation [of quantum mechanics], from which the conventional formulation can be *deduced*" (Everett 1957, 315). Here we will take Everett's proposal to regard the linear dynamics as a complete and accurate description of the time-evolution of every physical system seriously, perhaps more seriously than he did, and then examine its consequences.

One reason for taking Everett's proposal seriously is that there is some empirical support for believing that physical systems always evolve linearly. Whenever we have been able to perform the appropriate type of experiment, the systems we have observed have always evolved linearly. These experiments, however, are generally extremely difficult to perform. Roughly speaking, the object system must be one that can be well isolated from interactions with its environment, and the measurement interaction itself must be very precise. This means that we have only been able to perform the appropriate measurements on very simple systems that are easily isolated from their environments. Even so, our actual experiments have provided the basis for something of an inductive argument for the universal validity of the linear dynamics. It is an inductive argument on the complexity of a system and the difficulty in isolating it from its environment: we have empirical evidence that very simple easily isolated systems (fundamental particles) behave linearly, that slightly more complex and difficult to isolate systems (atoms) behave linearly, that slightly more complex and difficult to isolate systems (small molecules) behave linearly, etc. As we move to more complex and difficult to isolate systems, the experiments become more difficult to perform and relevant empirical evidence consequently becomes rarer. One might nonetheless take the evidence we do have as providing some empirical support for the proposal that all systems behave linearly.

But again, the collapse postulate apparently plays a critical role in the standard theory of quantum mechanics. Among other things, it helps to explain why we always get a determinate result to a measurement, why we get the same result when we repeat a measurement on an undisturbed system, why our measurement results are randomly distributed and have the relative frequencies that they do, and why there appear to be nonlocal quantum-mechanical effects. In order to take quantum mechanics without the collapse postulate seriously one must presumably find some way of providing such explanations without the collapse postulate. Our initial strategy is to drop the collapse postulate, then see how much we can do without it. It turns out that we can do more than one might at first imagine but less than one might reasonably want.

Quantum mechanics without the collapse postulate, what Albert calls the "bare" theory, has several properties that seem to be relevant to recapturing some of what the collapse postulate provides for the standard theory.[2] These properties tell us what an observer would report under specified physical circumstances. For example, if the bare theory is true, then we can show such things as the following: (1) After making a perfect measurement of any observable, an ideal measuring device M will be in an eigenstate of answering the question "Did you get a definite, unambiguous result?" with "Yes", (2) After a second perfect measurement of the same observable, M will be in an eigenstate of answering the question "Did you get the same result for both measurements?" with "Yes" if the object system is undisturbed between measurements, (3) If M measures the same observable of each of an infinite number of systems all in the same initial state, then M will approach an eigenstate of answering the question "Were your results randomly distributed with the same relative frequencies the standard theory of quantum mechanics would predict?" with "Yes" as the number of observations gets large, and (4) If an appropriate sequence of EPR experiments is performed on systems all in the same initial state, then the measuring devices involved will approach an eigenstate of answering the question "Were your results compatible with the Bell-type inequalities?" with "No" as the number of observations gets large.

While such properties are certainly suggestive, we presumably want more than these from a satisfactory physical theory. There will be more to say about this later, but it may help to have at least a vague idea of what is at stake from the beginning. According to the bare theory, an observer who begins in an eigenstate of being ready to make a measurement would end up in an eigenstate of reporting that he has an ordinary, determinate result to his measurement. This might mean that the observer believes that he has a determinate measurement resu t, but in the context of the bare theory this would not generally mean that there is any determinate result that the observer believes he has. Contrary to what Everett and others have claimed, the bare theory does not make the same empirical predictions as the standard theory; rather, the bare theory at best provides an explanation for why it might appear to an observer that the standard theory's empirical predictions are true when they are in fact false. That is, the bare theory provides the basis for claiming in some circumstances that some of one's beliefs are the result of an illusion. It cannot, however, generally do even this. Each of the bare theory's suggestive properties have the form 'If such-and-such physical conditions hold for an observer and his object system, then the observer would end up in an eigenstate of reporting such-and-such a belief', but if the bare theory is true, if the linear dynamics always correctly describes the time-evolution of all physical systems, then the antecedent conditions would virtually never be satisfied;

rather, the physical state would typically be a superposition of the antecedent conditions being satisfied and not being satisfied. Since the bare theory has little to say about what this would be like, it is presumably left with little to say about an observer's actual experience.

Given that one wants more from a satisfactory version of quantum mechanics than what the bare theory provides, one might want to exploit its suggestive properties to formulate a satisfactory no-collapse version of quantum mechanics. One strategy would be to supplement the bare theory with conditions that are as weak as possible yet allow one to recover much more of the predictive and explanatory power of the standard theory than the bare theory alone does. The consequences of adding the following two conditions to the bare theory suggest the possibility of a much richer theory that can exploit the suggestive properties in making predictions and providing explanations: (1) if the state of an observer M is an eigenstate of making some report, then M makes that report; and (2) an ideal observer's reports concerning his own mental states and the mental states of his ideal friends (given that he has asked them what they believe, etc.) are always true. It turns out, however, that these particular conditions are too strong – it is impossible to formulate any no-collapse version of quantum mechanics for which both of these are always satisfied.

2

Let's begin by examining several of the bare theory's suggestive properties in order to see precisely what they say and why they are true. While these hold for any physical observable, we will keep things simple by considering only spin observables of spin-1/2 systems.

Suppose that M is a perfect x-spin measuring device in the following sense: it is constructed so that its pointer variable becomes perfectly correlated with the x-spin of its object system S without disturbing the x-spin of S. If S is initially in an x-spin up eigenstate, then M reports that the result of its measurement is x-spin up and leaves S in the x-spin up state, and if S is initially in an x-spin down eigenstate, then M reports that the result of its measurement is x-spin down and leaves S in the x-spin down state. Now consider what happens when we take the linear dynamics to be a complete and accurate description of the time-evolution of every physical system.

It follows from how M has been constructed and from the linear dynamics that if the initial state $|\psi_0\rangle$ of $M + S$ is

$$|r\rangle_M(\alpha|\uparrow\rangle_S + \beta|\downarrow\rangle_S)$$

the state $|\psi_1\rangle$ after M's x-spin measurement will be

$$(2) \qquad \alpha|\uparrow\rangle_M|\uparrow\rangle_S + \beta|\downarrow\rangle_M|\downarrow\rangle_S.$$

Here M's state has become entangled with S's state. Furthermore, assuming that α and β are both non-zero, $|\psi_1\rangle$ is not an eigenstate of M reporting a particular determinate x-spin result; rather, M is in a superposition of reporting mutually contradictory results. It is, however, easy to show that $|\psi_1\rangle$ is an eigenstate of M reporting that it got some determinate x-spin result, either x-spin up or x-spin down.

Determinate Result

Asking M whether it got a determinate x-spin result amounts to measuring a physical observable of M. Let $\hat{D}$ be an observable such that eigenvalue $+1$ corresponds to a state where M has the disposition to report "I did get a determinate result to my x-spin measurement" and eigenvalue -1 corresponds to any orthogonal state.[3] Since $|\uparrow\rangle_M|\uparrow\rangle_S$ corresponds to a state where M has recorded $\uparrow$ for the outcome of its measurement, if M is operating correctly, then it will report that it obtained a determinate x-spin result when in this state; and since the same is true for $|\downarrow\rangle_M|\downarrow\rangle_S$, both of these are eigenstates of $\hat{D}$ with eigenvalue $+1$. That is,

$$\hat{D}|\uparrow\rangle_M|\uparrow\rangle_S = |\uparrow\rangle_M|\uparrow\rangle_S \quad \text{and} \quad \hat{D}|\downarrow\rangle_M|\downarrow\rangle_S = |\downarrow\rangle_M|\downarrow\rangle_S.$$

So

$$\begin{aligned}\hat{D}|\psi_1\rangle &= \hat{D}(\alpha|\uparrow\rangle_M|\uparrow\rangle_S + \beta|\downarrow\rangle_M|\downarrow\rangle_S) \\ &= \alpha\hat{D}|\uparrow\rangle_M|\uparrow\rangle_S + \beta\hat{D}|\downarrow\rangle_M|\downarrow\rangle_S \\ &= \alpha|\uparrow\rangle_M|\uparrow\rangle_S + \beta|\downarrow\rangle_M|\downarrow\rangle_S \\ &= |\psi_1\rangle.\end{aligned}$$

Consequently, $|\psi_1\rangle$ is an eigenvector of $\hat{D}$ with eigenvalue $+1$, so M has the disposition to report "I did get a determinate result to my x-spin measurement." Further, if M repeats its measurement, it will report that its second result agrees with the first even though, at least in the ordinary sense, it failed to get a determinate result to either measurement.

Repeatability

Suppose M makes a second x-spin measurement. Let M_1 and M_2 be the registers M uses to record the first and second measurement results. If S is undisturbed between measurements, the state $|\psi_2\rangle$ after the second measurement will be

$$\alpha|\uparrow\rangle_{M_1}|\uparrow\rangle_{M_2}|\uparrow\rangle_S + \beta|\downarrow\rangle_{M_1}|\downarrow\rangle_{M_2}|\downarrow\rangle_S$$

Let $\hat{C}$ be an observable such that states where M has the disposition to report that its first and second x-spin measurements agree (that is, they are both $\uparrow$ or both $\downarrow$) correspond to eigenvalue $+1$. Since registers M_1 and M_2 agree in the states represented by $|\uparrow\rangle_{M_1}|\uparrow\rangle_{M_2}|\uparrow\rangle_S$ and $|\downarrow\rangle_{M_1}|\downarrow\rangle_{M_2}|\downarrow\rangle_S$,, it follows that

$$\hat{C}|\uparrow\rangle_{M_1}|\uparrow\rangle_{M_2}|\uparrow\rangle_S = |\uparrow\rangle_{M_1}|\uparrow\rangle_{M_2}|\uparrow\rangle_S$$

and

$$\hat{C}|\downarrow\rangle_{M_1}|\downarrow\rangle_{M_2}|\downarrow\rangle_S = |\downarrow\rangle_{M_1}|\downarrow\rangle_{M_2}|\downarrow\rangle_S.$$

Since $|\psi_2\rangle$ is just a linear combination of eigenvectors of $\hat{C}$ with eigenvalue $+1$, it is also an eigenvector of $\hat{C}$ with eigenvalue $+1$, so M has the disposition to report that its first and second x-spin measurements agree.

Agreement

Suppose that rather than being interpreted as registers of the same measuring device, M_1 and M_2 are taken to be different measuring devices capable of comparing their results. It immediately follows from the repeatability property above that if M_1 and M_2 compare their results, they will report that they agree.

Given these properties, then, M will report that it got a determinate x-spin result ($\uparrow$ or $\downarrow$) regardless of whether S is initially in an x-spin eigenstate; if M remeasures S, it will report that its second result agrees with its first; and if a second measuring device measures the x-spin S and the two measuring devices compare results, both will report that their results agree. But again, if α and β are both non-zero, then none of these measurements in fact yield a determinate result, at least not in the ordinary sense. There are well-defined states resulting from M's interaction with S, but when M reports that it obtained a determinate result of $\uparrow$ or $\downarrow$ on its first measurement, for example, its report is *false* since $|\psi_1\rangle$ is not a state where M has recorded the result $\uparrow$ and it is not a state where M has recorded the result $\downarrow$, and these are the only possible determinate results here. While one might reasonably interpret M's report to mean that the physical state is either $|\uparrow\rangle_M|\uparrow\rangle_S$ or $|\downarrow\rangle_M|\downarrow\rangle_S$, while it may even *seem* to M that this is the case if M is a sentient observer, M is in fact in a superposition of having recorded mutually incompatible results. In order for M to correctly determine its state, it would have to measure an appropriate observable of the composite system $M + S$, but such a measurement would generally be very difficult to perform for reasons mentioned earlier.

Relative Frequency

Consider a system T consisting of a measuring device M and an infinite set of systems $S_1, S_2, S_3, \ldots, S_n \ldots$, each of which is initially in the state $\alpha|\uparrow\rangle_{S_n} + \beta|\downarrow\rangle_{S_n}$, where $|\uparrow\rangle_{S_n}$ and $|\downarrow\rangle_{S_n}$ are x-spin eigenstates and α and β are non-zero. Let T_n be the system consisting of M's first n registers and systems S_1 through S_n. The Hilbert space corresponding to T_n is

(7) $$\mathcal{H}_n \quad \mathcal{M}\otimes\mathcal{S}_1\otimes\cdots\otimes\mathcal{S}_{n-1}\otimes\mathcal{S}_n,$$

where $\mathcal{M}$ is the Hilbert space corresponding to the first n registers of the measuring device M, $\mathcal{S}_1$ is the Hilbert space corresponding to the system S_1, etc.

Suppose M makes an x-spin measurement on each S_n in turn. The states before and after the nth measurement might be represented by elements of $\mathcal{H}_n$. For example, the state of T_1 before the first measurement corresponds to the vector

(8) $$|r\rangle_M(\alpha|\uparrow\rangle_{S_1} + \beta|\downarrow\rangle_{S_1}),$$

which is an element of $\mathcal{H}_1$. After the first measurement, T_1 will be in the state corresponding to

$$\alpha|\uparrow\rangle_M|\uparrow\rangle_{S_1} + \beta|\downarrow\rangle_M|\downarrow\rangle_{S_1}$$

Similarly, the state of T_2 in $\mathcal{H}_2$ before the second measurement is the state corresponding to

$$\text{10)} \qquad \alpha|\uparrow, r\rangle_M \ \uparrow\rangle_{S_1} + \beta|\downarrow, r\rangle_M \ \downarrow\rangle_{S_1}(\alpha|\uparrow\rangle_{S_2} + \beta|\downarrow s_2\rangle)$$

and after the second measurement,

$$\begin{aligned} \alpha^2|\uparrow, \uparrow\rangle_M|\uparrow\rangle_{S_1}|\uparrow\rangle_{S_2} + \alpha\beta|\uparrow, \downarrow\rangle_M|\uparrow\rangle_{S_1}|\downarrow\rangle_{S_2} \\ + \beta\alpha|\downarrow, \uparrow\rangle_M \ \downarrow\rangle_{S_1}|\uparrow\rangle_{S_2} \\ + \beta^2|\downarrow, \downarrow\rangle_M|\downarrow\rangle_{S_1}|\downarrow\rangle_{S_2}. \end{aligned}$$

There are 2^n terms in the vector $|\psi_n\rangle$ representing the state of T_n after n measurements when written in this basis.

Let's call this the $\hat{X}_n$-basis, and let the observable $\hat{X}_n$ be such that T_n is in an eigenstate of $\hat{X}_n$ if and only if M is determinately reporting a particular sequence of results for the first n measurements. Given this, one can define a relative-frequency operator $\hat{F}_n(\epsilon)$ such that $\hat{F}_n(\epsilon)|\psi_n\rangle = |\psi_n\rangle$ if and only if $\hat{X}_n|\psi_n\rangle = \lambda|\psi_n\rangle$ and $(n-m)/n = |\alpha|^2 \pm \epsilon$, for $\epsilon > 0$, where $n - m$ is the number of $\uparrow$-results that M has after n measurements. If T_n is in an eigenstate of $\hat{F}_n(\epsilon)$, then M is in an eigenstate of displaying that the ratio of the number of $\uparrow$-results $n - m$ to the total number of results n is within ϵ of $|\alpha|^2$ – that is, that the ratio of $\uparrow$-results to $\downarrow$-results is about what the standard theory would predict. Note, however, that since $|\psi_n\rangle$ is not an eigenvector of $\hat{X}_n$ for any finite n, it is also not an eigenvector of $\hat{F}_n(\epsilon)$ for any finite n.

Write $|\psi_n\rangle$ in the X_n-basis, and let $|\chi(\epsilon)_n\rangle$ be the sum of those terms where the ratio of the number of $\uparrow$-results $n - m$ to the total number of results n is within ϵ of $|\alpha|^2$. In other words, let $|\chi(\epsilon)_n\rangle$ be the sum of the terms in the $\hat{X}_n$-expansion of $|\psi_n\rangle$ that *are* eigenvectors of $\hat{F}_n(\epsilon)$ with eigenvalue $+1$. Note that since $|\langle\chi(\epsilon)_n|\chi(\epsilon)_n\rangle|^2 < 1$ for all finite n, $|\chi(\epsilon)_n\rangle$ does not correspond to the state of any system; rather, $|\chi(\epsilon)_n\rangle$ is just the orthogonal projection of the state $|\psi_n\rangle$ onto the $\lambda = +1$ eigenspace of $\hat{F}_n(\epsilon)$. Finally, note that for all n, $|\chi(\epsilon)_n\rangle$ is an eigenvector of $\hat{F}_n(\epsilon)$ with eigenvalue $+1$.

Roughly speaking, the following lemma says that, for all $\epsilon > 0$, the magnitude of the component of $|\psi_n\rangle$ that is an eigenvector of $\hat{F}_n(\epsilon)$ with eigenvalue $+1$ goes to one as n gets large. Since the magnitude of $|\psi_n\rangle$ is one for all n, this will mean that $|\psi_n\rangle$ must approach an eigenvector of $\hat{F}_n(\epsilon)$ for all ϵ. We will then interpret this to mean that T_n approaches

$$\uparrow\rangle_{S_1} \uparrow_{S_2} \quad \downarrow\rangle_{S_{n-1}} \downarrow\rangle_{S_n}$$

$$_M | \downarrow\rangle_{S_1} | \downarrow\rangle_{S_2} \quad \downarrow\rangle_{S_n}$$

with 2^n such terms.

We will first partition these terms into equivalence classes by the relative frequency of $\uparrow$-results in each term, then consider a measure μ_n that assigns a real number to each equivalence class. For a given n, all terms where the number of $\uparrow$-results equals $n - m$ will be in the same equivalence class $\mathscr{B}_m$. For a given n, $\mu_n(\mathscr{B}_m)$ will be the sum of the squares of each of the coefficients of the terms in $\mathscr{B}_m$. For a given n, then, there will be n equivalence classes, and

$$\mu_n(\mathscr{B}_m) = \binom{n}{m} |\alpha^{n-m}\beta^m|^2, \tag{13}$$

which means that

$$\sum_m \binom{n}{m} |\alpha^{n-m}\beta^m|^2 = |\langle\chi(\epsilon)_n|\chi(\epsilon)_n\rangle|^2,$$

where the sum is over all m such that $m \leq n$ and $|\alpha|^2 - \epsilon \leq (n - m)/n \leq |\alpha^2 + \epsilon$ – that is, the sum is over all m where the ratio of the number of $\uparrow$-results to the total number of results is within ϵ of $|\alpha|^2$. Since it is a basic result of probability theory that

$$\lim_{n\to\infty} \sum_m \binom{n}{m} |\alpha^{n-m}\beta^m|^2 = 1$$

$|\langle\chi(\epsilon)_n|\chi(\epsilon)_n\rangle|^2 \to 1$ as $n \to \infty$ □

Given this, it is easy to show that, as n gets large, $|\psi_n\rangle$ gets arbitrarily close to an eigenvector of $\hat{F}_n(\epsilon)$ with eigenvalue $+1$.

THEOREM 1. *For all* $\epsilon > 0$, $\lim_{n\to\infty} |\langle\psi_n - \chi(\epsilon)_n|\psi_n - \chi(\epsilon)_n\rangle|^2 = 0$.

Proof. Since $|\chi(\epsilon)_n\rangle$ is an orthogonal projection of $|\psi_n\rangle$ onto a subspace of $\mathcal{H}_n$, namely the eigenspace corresponding to eigenvalue $+1$ of $\hat{F}_n(\epsilon)$, $|\psi_n - \chi(\epsilon)_n\rangle$ is orthogonal to $|\chi(\epsilon)_n\rangle$. So, by the Pythagorean theorem,

$$|\langle\psi_n - \chi(\epsilon)_n|\psi_n - \chi(\epsilon)_n\rangle|^2 = |\langle\psi_n|\psi_n\rangle|^2 - |\langle\chi(\epsilon)_n|\chi(\epsilon)_n|^2.$$

Since $|\langle\psi_n|\psi_n\rangle|^2 = 1$ for all n and since $|\langle\chi(\epsilon)_n|\chi(\epsilon)_n\rangle|^2 \to 1$ as $n \to \infty$, $|\langle\psi_n - \chi(\epsilon)_n|\psi_n - \chi(\epsilon)_n\rangle|^2 \to 0$ as $n \to \infty$. □

It follows that if M makes measurements of the x-spin on each system, then it will approach an eigenstate of answering the question "Were your results distributed with the same frequencies the standard theory of quantum mechanics would predict?" with "Yes" as the number of observations gets large. Or more specifically, M will approach an eigenstate of reporting that

$$\lim_{n\to\infty} \frac{n-m}{n} = |\alpha|^2$$

where $n - m$ is the number of $\uparrow$-results and n is the total number of x-spin results. Note that while M approaches an eigenstate of reporting that it has a determinate sequence of results that are distributed just as predicted by ordinary quantum mechanics, it does not approach a state where it actually has a determinate set of results.

Randomness

As the number of x-spin measurements gets large, T_n also approaches a state where M would answer the question "Were your results randomly distributed?" with "Yes". Given the argument for the relative-frequency property, this is fairly straightforward. Write $|\psi_n\rangle$, the state of T_n after n measurements, in the $\hat{X}_n$-basis. Each term in the expansion will correspond to a different sequence of measurement results, and each sequence of length n will be represented by some term. For all n, the norm squared of the coefficient associated with each sequence determines a measure μ_n on the set of all length n sequences Q_n such that $\mu_n(Q_n) = 1$ and such that if the μ_n-measure on the subset of random sequences of Q_n goes to one as n gets large, then M would approach an eigenstate of reporting that its x-spin results were randomly distributed in the limit. So M will approach an eigenstate of reporting that its results were randomly distributed for any notion of randomness where the μ_n-measure of random sequences of length n goes to one as n gets large. Presumably any satisfactory notion of randomness would have this property. Note that since the bare theory is in fact perfectly deterministic, there is at most the *appearance* of randomness.

The relative-frequency and randomness properties have the following corollary: if any single experiment whatsoever yields a state where the amplitude of M recording the result r_i is α_i, then M will approach a state where it reports that the result r_i is randomly distributed with relative frequency $|\alpha_i|^2$ in the limit as an infinite number of identical experiments are performed. Let's call this the *general limiting property*.[4] Two concrete examples will help to show how general it is.

In these examples we will use the following three noncommuting spin observables of the system S: x-spin, which has eigenstates $|\uparrow_x\rangle_S$ and $|\downarrow_x\rangle_S$; z-spin, which has eigenstates $|\uparrow_z\rangle_S = 1/\sqrt{2}(|\downarrow_x\rangle_S + |\uparrow_x\rangle_S)$ and $|\downarrow_z\rangle_S = 1/\sqrt{2}(|\downarrow_x\rangle_S - |\uparrow_x\rangle_S)$; and u-spin, which has eigenstates $|\uparrow_u\rangle_S = \sqrt{3}/2|\uparrow_x\rangle_S - 1/2|\downarrow_x\rangle_S$ and $|\downarrow_u\rangle_S = 1/2|\uparrow_x\rangle_S + \sqrt{3}/2|\downarrow_x\rangle_S$.[5]

Suppose that S is initially in the state $|\uparrow_z\rangle_S$. Observer A measures x-spin, then observer B measures u-spin. What is the probability for each of the possible outcomes of the two measurements? The standard theory of quantum mechanics tells us that when A measures the x-spin of S, S will nonlinearly collapse to the state $|\uparrow_x\rangle_S$ with probability 1/2 and collapse to the state $|\downarrow_x\rangle_S$ with probability 1/2 – that is, $p(\uparrow_x) = p(\downarrow_x) = 1/2$. When B then measures the u-spin of S, S will similarly collapse to an eigenstate of u-spin. If S initially collapsed to $|\uparrow_x\rangle_S$, then the probability of $|\uparrow_u\rangle_S$ is 3/4 and the probability of $|\downarrow_u\rangle_S$ is 1/4, but if S initially collapsed to $|\downarrow_x\rangle_S$, then the probability of $|\uparrow_u\rangle_S$ is 1/4 and the probability of $|\downarrow_u\rangle_S$ is 3/4. It follows that $p(\uparrow_x \text{ and } \uparrow_u) = 3/8$, $p(\uparrow_x \text{ and } \downarrow_u) = 1/8$, $p(\downarrow_x \text{ and } \uparrow_u) = 1/8$, and $p(\downarrow_x \text{ and } \downarrow_u) = 3/8$. Note that these joint probabilities are calculated in the standard theory by supposing that S *collapses* to an eigenstate of x-spin on the first measurement.

Now suppose that the linear dynamics correctly describes the evolution of the systems – what relative frequencies will A and B report in the limit as they perform an infinite number of such measurements? The initial state is

$$|r\rangle_B\,|r\rangle_A\,\frac{1}{\sqrt{2}}(|\uparrow_x\rangle_S + |\downarrow_x\rangle_S).$$

The linear dynamics tells us that after the x-spin measurement the state will be

$$|r\rangle_B\,\frac{1}{\sqrt{2}}(|\uparrow_x\rangle_A\,|\uparrow_x\rangle_S + |\downarrow_x\rangle_A\,|\downarrow_x\rangle_S),$$

where $|\uparrow_x\rangle_A$ represents a state where measuring device A reports x-spin up, etc.

Similarly, after the u-spin measurement that state will be

(20)
$$\frac{\sqrt{3}}{2\sqrt{2}}|\uparrow_u\rangle_B\,|\uparrow_x\rangle_A\,|\uparrow_u\rangle_S + \frac{1}{2\sqrt{2}}|\downarrow_u\rangle_B|\uparrow_x\rangle_A\,|\downarrow_u\rangle_S)$$
$$-\frac{1}{2\sqrt{2}}|\uparrow_u\rangle_B\,|\downarrow_x\rangle_A\,|\uparrow_u\rangle_S) + \frac{\sqrt{3}}{2\sqrt{2}}|\downarrow_u\rangle_B\,|\downarrow_x\rangle_A\,|\downarrow_u\rangle_S).$$

Consequently, it follows from the bare theory's general limiting property that A and B will approach a state where they report that the possible outcomes were randomly distributed with the same relative frequencies predicted by the standard theory: $p(\uparrow_u \text{ and } \uparrow_x) = [\sqrt{3}/(2\sqrt{2})]^2 = 3/8$, etc.[6]

For another example suppose that two systems S_A and S_B are initially in the EPR state

(21)
$$\frac{1}{\sqrt{2}}(|\downarrow_x\rangle_{S_A}\,|\uparrow_x\rangle_{S_B} - |\uparrow_x\rangle_{S_A}\,|\downarrow_x\rangle_{S_B})$$

and that A and B make space-like separate measurements of their respective systems. When A measures the x-spin of S_A, the standard theory predicts that the composite system will collapse to the state $|\downarrow_x\rangle_{S_A}\,|\uparrow_x\rangle_{S_B}$ with probability 1/2 and collapse to the state $|\uparrow_x\rangle_{S_A}\,|\downarrow_x\rangle_{S_B}$ with probability 1/2. This means that $p(\uparrow_x @ A) = p(\downarrow_x @ A) = 1/2$. If the composite system collapses to $|\downarrow_x\rangle_{S_A}\,|\uparrow_x\rangle_{S_B}$, then $p(\uparrow_u @ B) = 3/4$ and $p(\downarrow_u @ B) = 1/4$. If the composite system collapses to $|\uparrow_x\rangle_{S_A}\,|\downarrow_x\rangle_{S_B}$, then $p(\uparrow_u @ B) = 1/4$ and $p(\downarrow_u @ B) = 3/4$. So the standard theory predicts that $p(\uparrow_x @ A \text{ and } \uparrow_u @ B) =$

1/8, $p(\uparrow_x @ A$ and $\downarrow_u @ B) = 3/8$, $p(\downarrow_x @ A$ and $\uparrow_u @ B) = 3/8$, and $p(\downarrow_x @ A$ and $\downarrow_u @ B) = 1/8$.

What does the bare theory predict in the limit as this experiment is performed an infinite number of times? After A's x-spin measurement and B's u-spin measurement, the linear dynamics tells us that the state of the composite system will be

(22) $$\frac{1}{\;}|\uparrow_u\rangle_B|\uparrow_x\rangle_A|\uparrow_u\rangle_{S_B}|\uparrow_x\rangle_{S_A}$$

$$+\frac{\sqrt{3}}{2\sqrt{2}} \qquad |\uparrow_u\rangle_{S_B}|\downarrow_x\rangle_{S_A}$$

$$\frac{\sqrt{3}}{2\sqrt{2}} \qquad |\uparrow_x\rangle_A|\downarrow_u\rangle_{S_B}|\uparrow_x\rangle_{S_A}$$

$$+\frac{1}{2\sqrt{2}}|\downarrow_u\rangle_B|\downarrow_x\rangle_A|\downarrow_u\rangle_{S_B}|\downarrow_x\rangle_{S_A}$$

So given the general limiting property, A and B will approach an eigenstate of reporting that their measurement results were randomly distributed and statistically correlated in just the way the standard theory predicts: $p(\uparrow_x @ A$ and $\uparrow_u @ B) = [-1/(2\sqrt{2})]^2 = 1/8$, etc. Since in the limit A and B will similarly agree that they got the statistical correlations predicted by standard theory for any other pair of spin observables and since the standard theory's predictions fail to satisfy the Bell-type inequalities, if they perform an appropriate sequence of different experiments, then they will approach an eigenstate of reporting that their results fail to satisfy the Bell-type inequalities. Since what happens at A seems to instantaneously influence the result of a measurement at B and the other way around, the observers may be tempted to conclude that there is a nonlocal causal connection between their measurements. Note, however, that since the linear dynamics can be written in a perfectly local form, there are in fact no nonlocal causal connections in the bare theory. One might explain that since A and B in fact have no measurement results, there are in fact no determinate events that are correlated, so it is not at all surprising that there are no nonlocal causal connections between events at A and events at B. Just as reports of determinate results, relative frequencies, and randomness would generally be explained by the bare theory as illusions, the apparent nonlocality here would be just that, apparent.

3.

Since the bare theory's suggestive properties suggest the possibility of accounting for an observer's reports concerning his own experience as illusions generated by the linear dynamics, one might be tempted to take the bare theory to be a complete and accurate physical theory. There is something perversely attractive in this, but it is very difficult to take the bare theory to be the whole story. As Albert has argued, for example, if the bare theory is true, then there will be matters of fact concerning what we think about the relative frequencies of our measurement results *only* in the limit as the number of those measurements goes to infinity: ". . . if the bare theory is true . . . , then there can't now be any matter of fact (not withstanding our delusion that there is one) about what we take those frequencies to *be*." Further,

> If the bare theory is true, then it seems extraordinarily unlikely that the present quantum state of the world can possibly be one of those in which there's even a matter of fact about whether or not any sentient experimenters exist at all. And of course in the event that there *isn't* any matter of fact about whether or not any sentient experimenters exist, then it becomes unintelligible to inquire . . . about what sorts of things experimenters will *report* (Albert 1992, 124–5).

The point again is that the suggestive properties at best only tell us what a good observer would believe about his experiences in certain specific circumstances, circumstances that the linear dynamics itself tells us would rarely if ever obtain.

Given the suggestive properties, then, perhaps the most difficult question is what to do with them. They are too suggestive to ignore, so we ought to exploit them some way. After all, they show us that much of what the collapse postulate is supposed to do for the standard theory can be done in certain circumstances with just the linear dynamics. The problem is how to best exploit these properties. One would presumably want to add something that would at least ensure that there are ordinary matters of fact concerning the mental states of sentient observers.

As a starting point, let's suppose that the linear dynamics is a complete and accurate description of the time-evolution of the physical state and that the following two conditions are satisfied:

> *Eigenstate condition:* If the state of an observer M is an eigenstate of making some report, then M makes that report.
>
> *Reliability condition:* An ideal observer's reports concerning his own mental states and the mental states of his ideal friends (given that they have told him what their mental states are, etc.) are always true.

The eigenstate condition is so basic to our understanding of quantum-mechanical states that it is usually assumed without comment. The reliability condition is also plausible. It may be that real observers fall short of the ideal, but one would presumably be surprised to find that it is physically impossible for there to be an ideal observer in this sense. If it were impossible, then first-person authority would be limited by our physical theory.

Given the eigenstate condition, the bare theory's suggestive properties tell us what an observer would report in certain circumstances. So the reliability condition directly determines what some of the observer's beliefs would be in those circumstances. But what about those beliefs that are not directly determined by the fact that the observer is in an eigenstate of making a particular report? One answer is that the reliability condition together with those beliefs that are directly determined by the observer's physical state *indirectly* determine many of those beliefs that are not.

By assuming that the linear dynamics is universally true and by imposing the two conditions above, one is forced to conclude that an observer's *mental state* is not fully determined by his *physical state*. One consequence of this is that an observer might generally have a determinate belief concerning what his experience has been even when his physical state is a superposition of states corresponding to different beliefs. Suppose an observer makes a perfect x-spin measurement of an object system initially in an eigenstate of z-spin and ends up in the physical state predicted by the linear dynamics. The determinate result property tells us that the observer will report that he got either x-spin up or x-spin down as the result of his measurement. If we assume that the observer's reports concerning his own mental state are true, then the determinate result property requires his mental state to correspond to one or the other of the two possible x-spin results – that is, he must believe that the result was $\uparrow$ or he must believe that the result was $\downarrow$. But which belief will he in fact have? The relative-frequency and randomness properties help here. They require that the observer's physical state approach an eigenstate state of reporting that his measurement results were randomly distributed with the same relative frequencies predicted by the standard theory in the limit. In order for this report to be true, the observer's mental state would have to approach a state where his beliefs concerning the x-spin results of his measurements really were randomly distributed with the appropriate relative frequencies. If we suppose that the observer's mental dynamics is trial-independent, that the rules that determine which x-spin belief he will end up with do not change from measurement to measurement, each x-spin belief would have to be randomly determined by probabilities equal to the limiting relative frequencies. This means that the result of a single x-spin measurement of a system initially in an eigenstate of z-spin would with probability 1/2 be x-spin up and with probability 1/2 be x-spin down, which is just what we want.

4.

While the eigenstate and reliability conditions may serve as a convenient starting point for exploiting the bare theory's suggestive properties and while they might at first seem plausible, no no-collapse formulation of quantum mechanics can generally satisfy both of these conditions. In the limit as an observer M performs an infinite number of x-spin measurements, for example, M will approach an eigenstate of reporting that there exists a determinate sequence of results that correctly describes his experience. But for each possible sequence of results, M will also approach an eigenstate of denying that that particular sequence correctly describes his experience. This is because the norm squared of the coefficient on each term of the state after n measurements $|\psi_n\rangle$ written in the $\hat{X}_n$-basis goes to zero as n gets large. Given this, either the eigenstate condition or one of M's reports must be false.

The following is perhaps a more vivid example of what can happen if one assumes both the eigenstate and reliability conditions in a no-collapse formulation of quantum mechanics. Suppose M measures the x-spin of each of an infinite sequence of systems all initially in eigenstates of z-spin. By the eigenstate condition, M will report that he got a determinate result to each measurement – either x-spin up or x-spin down. The reliability condition consequently requires there to be some sequence of results that M believes that he got. Suppose that after each measurement M decides whether or not to carefully *remeasure* the x-spin of the system. Suppose further that an infinite number of M's first-measurement results are x-spin up. presumably this would happen with probability one in an empirically adequate formulation of quantum mechanics, and suppose that by pure chance, say, M decides to remeasure exactly those systems that he got x-spin up for on his first measurement. By the eigenstate condition, M will report that he got the same result for both measurements whenever he remeasures a system. So by the reliability condition and the assumption that M remeasured only those systems where he got x-spin up for his first result, *all of his second measurements must also yield the result x-spin up*. On the other hand, the eigenstate condition together with the relative-frequency and randomness properties entail that M will approach a state where he reports that the second-measurement results were randomly distributed with half of the results x-spin up. By the reliability condition, then, *an infinite number of M's second measurements must fail to yield the result x-spin up*. Since it is impossible for an infinite number of results to be both x-spin up and not x-spin up, the eigenstate and reliability conditions cannot both be true here.

One might want to weaken one or the other of the two conditions. The eigenstate condition, for example, might be changed to say that if the state of an observer M is an eigenstate of making some report, then with *probability one M* makes that report. Or one might be willing to accept the conclusion that no observer can generally make reliable reports concerning his own experiences. In any case, the moral here is that any no-collapse theory that can make coherent empirical predictions in the limit must violate either the eigenstate condition or the reliability condition or both.

5

Quantum mechanics without the collapse postulate has several suggestive properties. These properties tell us what an observer would report in various situations. If he measures a system that is not in an eigenstate of the observable being measured, then the linear dynamics together with what it means to be a good observer tells us that he will nonetheless report that he got a ordinary, determinate report. If he carefully repeats the measurement, then he will report that he got the same result for the second measurement. If he and his friends carefully measure the state of the same system, then they will report that their results agree. Further, as an observer measures the same observable of an infinite sequence of identically prepared systems, he will approach an eigenstate of reporting that the results were randomly distributed with the usual relative frequencies. Or stated another way, if any single experiment whatsoever yields a state where the amplitude of the observer recording the result r_i is α_i, then the observer would approach an eigenstate of reporting that the result r_i was randomly distributed with relative frequency $|\alpha_i|^2$ in the limit as an infinite number of identical experiments were performed. Among other things, this means that all of the joint probabilities come out right in the limit: if, for example, an appropriate sequence of EPR experiments were performed, the observers involved would approach an eigenstate of reporting that their empirical results violate the Bell-type inequalities in just the way the standard theory does.

The bare theory, however, does not provide what one might reasonably expect from a satisfactory physical theory. It makes determinate empirical predictions whenever one's object system happens to be in an eigenstate of the observable being measured, but if the linear dynamics is correct, this would virtually never happen. If the object system is not in an eigenstate of the observable being measured, then the bare theory provides a basis for explaining why an observer might believe that he has a determinate belief concerning the result when he in fact does not. But even this requires the observer to begin in an eigenstate of being ready to perform a reliable measurement on the particular object system, and if the linear dynamics is correct, this would virtually never happen. The bare theory's general limiting property tells us what an observer would report in the limit as an infinite number of identical experiments were performed, but whether or not the bare theory is correct, observers presumably never do anything like perform an infinite number of measurements on identically prepared systems. In short, the bare theory seems to have little to say that is directly relevant to the experiences of a real observer.

If one adds something like the eigenstate and reliability conditions to the bare theory, then the suggestive properties constrain an observer's beliefs in a way that suggests the possibility of a theory that exploits the suggestive properties to make empirical predictions which are much richer than those of the bare theory. While these conditions may sound plausible, however, it is impossible for any no-collapse theory to satisfy them.

NOTES

[1] See von Neumann (1932) for an early example of this formulation of quantum mechanics.
[2] Some of these properties were strongly suggested by Everett (1957). There have been several subsequent attempts to clarify the properties discussed by Everett and to determine their significance: see Hartle (1968), DeWitt (1971), Everett (1973), Graham (1973), and Albert (1992) for examples.
[3] Note that $\hat{D}$ is not the identity operator. For example, $\hat{D}|r\rangle_M|\uparrow\rangle_S \neq |r\rangle_M|\uparrow\rangle_S$.
[4] The general limiting property holds even if we allow for certain sorts of imperfect measurements. It can be shown, for example, that it remains true even if M measures a slightly different observable in each experiment as long as these observables are distributed about a mean observable in the limit.
[5] In order to work though these examples the following identities may be useful: $|\uparrow_x\rangle_S = \sqrt{3}/2|\uparrow_u\rangle_S + 1/2|\downarrow_u\rangle_S$ and $|\downarrow_x\rangle_S = -1/2|\uparrow_u\rangle_S + \sqrt{3}/2|\downarrow_u\rangle_S$.
[6] This state is analogous to the state of $M + S_1$ after one measurement in the argument for the relative-frequency property. Note that in the argument for the relative-frequency property, the relative frequencies that M would report in the limit were shown to be equal to the squares of the coefficients on the terms corresponding to each possible result after the first measurement.

REFERENCES

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

DeWitt, B. S.: 1971, 'The Many-Universes Interpretation of Quantum Mechanics', in *Foundations of Quantum Mechanics*, Academic Press, New York; reprinted in DeWitt and Graham (1973).

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

Everett, H., III: 1957, '"Relative State" Formulation of Quantum Mechanics', *Reviews of Modern Physics* **29**, 454–462; reprinted in DeWitt and Graham (1973).

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

Graham, N.: 1973, 'The Measurement of Relative Frequency', in DeWitt and Graham (1973).

Hartle, J. B.: 1968, 'Quantum Mechanics of Individual Systems', in *American Journal of Physics* **36**, 704–712.

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

Manuscript submitted May 27, 1994

Department of Philosophy
University of California
Irvine, CA 92717-4555
U.S.A.