Wanted Dead or Alive: Two Attempts to Solve Schrödinger's Paradox[1]

David Albert and Barry Loewer

Columbia University and Rutgers University

In a discussion of Schroedinger's views on quantum theory John Bell says that Schroedinger did not see how "to account for particle tracks in track chambers...and more generally for the definiteness, the particularity, of experience, as compared with the indefiniteness, the waviness, of the wave function. It is the problem he had had with his cat. He thought it could not be both dead and alive. But the wave function showed no such commitment, superposing the possibilities. Either the wave function as given by the Schroedinger equation is not everything or it is not right" (Bell 1987). At a recent conference Bell sermonized against the employment of "for all practical purpose" reasoning- he called it FAP reasoning-to solve this problem (Bell 1989). He argued that we should not be satisfied with any alleged solution which works "for all practical purposes" only while leaving conceptual puzzles unresolved. In particular we should not be satisfied with a cat which is alive (or dead) for all practical purposes only. In this paper we will be discussing two recent attempts to make quantum theory compatible with the definiteness of experience; attempts to guarantee that cats always end up either dead or alive. The first way, which is based on a proposal by Kochen, (Kochen 1986) and developed in different ways by Healey (Healey 1989), Dieks (Dieks 1989), and van Frassen (van Frassen 1981 and 1990) says that the wave function is not everything. The second is a recent theory of the collapse of the wave function due to Ghirardhi, Rimini, and Weber (GRW 1986 and Bell 1987) which says that Schroedinger's equation is not right. The first, we will argue, resorts to FAP thinking at a crucial point and for this reason turns out to be unsatisfactory. The second also seems to involve FAP thinking but in this case such thinking may be avoidable. So our conclusions concerning it are more cautious and tentative. We begin with a brief review of the reasoning that led Schroedinger to the problem he had with his cat.

Schroedinger imagined an arrangement like the following. A cat (whose initial state is $|R\rangle$) interacts with an electron so that if the electron's x-spin is up then the cat is killed and if the electron's x-spin is down then the cat remains alive. That is, at the start of the interaction if the state of the electron+cat is $|\uparrow\rangle|R\rangle$ ($|\downarrow\rangle|R\rangle$) then at its conclusion it will be $|\uparrow\rangle|Dead\rangle$ ($|\downarrow\rangle|Alive\rangle$). Quantum theory assures us that there exist Hamiltonians which characterize such interactions as this one. Further, it follows from the linearity of Schroedinger's dynamical equation that at the conclusion of a measurement of an electron whose initial state is $c_1|\uparrow\rangle + c_2|\downarrow\rangle$ the final state will be

PSA 1990, Volume 1, pp. 277-285

CAT $c_1|\uparrow>|Dead> + c_2|\downarrow>|Alive>$

As Bell says, this state superposes the possibilities of Dead and Alive. If one assumes, as standard interpretations of quantum theory do, that an observable has a well defined value for a system S when and only when S's quantum state is an eigenstate of that observable then the unfortunate cat who finds itself in CAT is neither dead nor alive. This struck Schroedinger (and Bell) and most others who have thought about it as intolerable.

Schroedinger's cat is a measuring device whose state of aliveness registers the spin of the electron. A similar consequence follows for any accurate measurement of an observable when the system being measured is not in an eigenstate of that observable. A perfectly accurate and non-destructive measurement of observable A of S by B of M is required to satisfy the condition:

> If A is an observable pertaining to system S which has values a_i and B is an observable pertaining to a measuring device M with values b_i-these are the measuring device's pointer positions in which the outcome is recorded- then an interaction between M and S is a measurement of A on S by B on M only if the Hamiltonian which characterizes the interaction guarantees that whenever S+M starts in state $|A=a_k>|R>$ it ends up in state $|a=A_k>|B=b_k>$ ($|R>$ is M's "ready to measure" state.)

The problem is that if S starts in a state which is not an eigenstate of A and if Schroedinger's equation characterizes the interaction then S+M will end up in the state

MEAS $SUMc_i|A=a_i>|B=b_i>$

which is not an eigenstate of AxB. If observables possess values only in their eigenstates then at the end of the measurement neither A nor B will possess a definite value. The problem of avoiding this intolerable conclusion has come to be called "the measurement problem."

There is a standard resolution of the measurement problem due to von Neuman (von Neuman 1955). He takes the second of Bell's options denying that Schroedinger's equation is always right. In the cat interaction the final state of cat+e is not CAT but is one of the states $|Dead>|\uparrow>$ or $|Alive>|\downarrow>$ with probabilities $c_1{}^2$ and $c_2{}^2$. More generally, the final state of S+M in a genuine measurement is not MEAS but one of the states $|A=a_k>|B=b_k$. This is called the "projection" or "collapse" postulate and it, according to von Neuman, and not Schroedinger's equation governs measurement interactions.

This is a very ingenious solution since it allows one to maintain that observables possess values only for systems in eigenstates. But it is very difficult to believe that the collapse described by von Neumann can be a real physical process. The main problem is that there are in principle experimentally detectable differences between the states which result from Schroedinger's equation and those which result from collapses and that for all those cases which have been experimentally investigated the final state has been found to be the one predicted by Schroedinger's equation. It is true that these experiments have always involved interactions between micro-systems (systems containing a few particles). In view of this, a proponent of the standard view can say (and many have) that genuine measurements must involve macroscopic devices. Their idea is that as long as the interaction is microscopic Schroedinger's equation obtains, but that when it involves a macroscopic system the collapse postulate obtains. But there seems to be no non-arbitrary way of drawing a line between micro and macro systems. At what point of complexity does an interaction which establishes a correlation between the observables of two systems become a measurement? And while the transition from microscopic to macroscopic is gradual the differences between the two kinds of dynamics are dramatic. The Schroedinger evolution is continuous, deterministic, and time reversible while the collapse dynamics is discontinuous, probabilistic, and time irreversible. It is simply incredible that at some point of complexity the linear dynamics are suspended and the collapse dynamics takes over. So those who endeavor to understand quantum theory as a true description of reality have looked for ways of avoiding von Neuman type collapses. There are two relatively new interpretations on the block which attempt to do just that.

1. The Modal Interpretation

The first "solution" which we will examine involves the idea that an observable of a system may have a value even when the state of that system is not an eigenstate of that observable. Different (though related) proposals of this sort have been made by Simon Kochen (1985), Dennis Dieks (1988), Richard Healey (1989) and Bas van Frassen (1981, 1990). Van Frassen calls his account "the modal interpretation" (1981) and we will adopt this name to cover all these proposals. The modal interpretation responds to the measurement problem by dropping the assumption that an observable pertaining to a system possesses a value only when the state of that system is an eigenstate of that observable. This is also what is done by certain so-called "hidden variable" theories (Bohm 1952). A hidden variable theory assumes that every dynamical variable possesses a precise value in every state. It turns out that if such theories are to reproduce the statistical predictions of quantum theory then they must exhibit some odd features. Specifically they must be non local and contextual (Bell 1987). The first means, roughly, that changes in a dynamical quantity that pertains to one component of a system may instantaneously affect the value of a dynamical quantity that refers to another component even if these components are spatially separated. The second means that precisely how a quantity is measured may affect the outcome of the measurement. These odd features have been taken as reasons for rejecting such theories, although these so called "no hidden variables proofs" are, in our view, far from conclusive (Albert and Loewer 1989).

The modal interpretation is seemingly more sophisticated. It allows only some of the observables which pertain to a subsystem of a larger system to possess values when the system is not in an eigenstate of that observable. This is done in such a way, it is hoped, as to supply outcomes to measurements and definite positions to pointers and definite states of aliveness to cats and so on. In describing the modal interpretation it will be useful to imagine a measurement of x-spin in which the result is recorded in the position of a pointer. The pointer points to UP if the spin is up and to DOWN if the spin is down (we will use |U> and |D> for the associated pointer states.) Recall that the measurement problem arises because the linear Schroedinger equation predicts that the post-measurement state of M+S is CAT (with the |U> replacing |Alive> etc.) and that state is not an eigenstate of the pointer's position. If we assume, as the standard interpretation of quantum theory does, that the pointer possesses a position only when its state is an eigenstate of position, then CAT is a state in which the pointer fails to have a position. Here is what the Healey-Dieks version of the modal interpretation says about this. The state CAT can be represented in various bases (as the sum of various complete sets of orthogonal vectors) but the particular representation used above has a special feature. As long as $c_1 \neq c_2$ it is the unique representation of the form SUM $c_i|a>_i|b>_i$ where the $|a>_i$ are orthogonal and the $|b>_i$ are orthogonal. Such representations are called "bi-orthonormal." Their interpretation is summarized by the following rules:

R1 If S is composed of S_1 and S_2 and the state of S is \$ with biorthonormal representation = $SUMc_i|a_i>_{S1}|b_i>_{S2}$ and the $|a>_i$ are eigestates of observable A and the $|b>_i$ are eigenstates of observable B then A is defined (has a value) on system S_1 and B is defined on S_2.

R2 If S is in the state referred to in R1 the probability that A has value a_i and B has value b_i is c_i^2.

At first sight it looks as though this proposal solves the measurement problem very nicely. In consequence of R1 the pointer position observable will possess a definite value at the conclusion of a measurement and the cat will end up either dead or alive. Furthermore, this is done without there being any collapse of state (the Schroedinger equation is the only description of state evolution). We might also add that this account does not run afoul of the no hidden-variable theorems (since not all observables obtain values via R1 and R2) and, while it has an inevitable non-locality the advocates of this view claim that it is no more objectionable than it is in standard quantum theory. There is a problem about what to say when there does not exist a unique bi-orthonormal decomposition (the usual presentations of Schroedinger's paradox involve just such a state) but perhaps this shouldn't be a big worry since almost every state (i.e. the measure on such states in the Hilbert space is 1) possesses a unique bi-orthonormal decomposition.

Van Frassen's (1990) version of the modal interpretation is similar to the Dieks-Healey version in that it allows observables to possess values for states which are not eigenstates of the observable but it works a bit differently. Suppose that M and S are systems whose interaction is governed by a Hamiltonian H. Then there is at most one maximal observable A of S and one observable B of M such that when S+M begins in state $|A=a_i>|R>$ it ends up in state $|A=a_i>|B=b_i>$. That is there are at most one pair of observables satisfying the conditions on a non-disturbing and perfectly accurate measurement. Van Frassen says that in an interaction governed by H the observables A and B are well defined on S and M respectively at the conclusion of the measurement. Van Frassen's account looks not at the final state (as Dieks-Healey does) but rather at the Hamiltonian which governs the interaction which produces the final state. This results in at least one advantage since it specifies that A and B possess definite values even when the post measurement state is one in which some of the c_i are equal. We will focus most of our discussion on the Dieks-Healey version although many of our remarks apply to both modal interpretations.

There are two respects in which the modal interpretation as so far described is incomplete. The first is that the prescription for determining the values of defined observables on a subsystem is underspecified. A many particle system can be decomposed many ways into subsystems with each way possessing a biorthonormal state representation. Rules need to be introduced to guarantee that the values of observables which are defined according to one such decomposition do not conflict with values defined in another. It is likely that this can be done in more than one way and so whatever rules are chosen may appear to be arbitrary. The second, and more serious, lacuna concerns the dynamics of the physical quantities defined in accord with (R1). As the state \$(t) of a composite system evolves in accordance with Schroedinger's equation the vectors in the biorthonormal representation of \$(t) will typically change. In consequence the observables defined on the system's subsystems in accord with (R1) will change. The question that the dynamics needs to answer is how the values of all those observables that have values at t+ are related to the values of those observables that have values at t.

We now come to an objection which is, in our view, fatal to all versions of the modal interpretation. The problem is that measurements which are actually performed almost never satisfy the condition we placed on measurements. In a real measurement there is always some probability of the measuring device making an error. For this reason in a realistic measurement the post-measurement state of M+e will be

$$\text{CAT*} \quad c_3|\uparrow>|U>+c_4|\uparrow>|D> + c_5|\downarrow>|U> + c_6|\downarrow>|D>$$

The components $|\downarrow>|U>$ and $|\uparrow>|D>$ represent errors. If the measurement is a good one then the amplitudes of these error components will be small but, as we have noted, in realistic measurements they will inevitably be non-zero. CAT* is obviously not a biorthonormal representation. The biorthonormal representation of this state is

$$\text{BCAT*} \quad c_7|\uparrow_{@}>|G> + c_8|\downarrow_{@}>|H>$$

where $|\uparrow_{@}>$ represents the state in which spin in direction @ (not x-spin) is up and $|G>$ represents a state of the pointer. (R1) does not assign a definite x-spin to the electron in BCAT* although it does assign a definite value to the spin in the @ direction. This is a bit disconcerting since M is supposed to be measuring x-spin. But the real problem is that $|G>$ and $|H>$ are not eigenstates of the observable corresponding to the pointer's position. That is, (R1) fails to assign a definite position to the pointer. If all M's observables which possess values are determined by (R1) then, according to the Dieks-Healey account, there simply isn't going to be a matter of fact when CAT* obtains about whether the pointer is pointing to `up' or to `down'! (And similarly, if there is some small chance of the cat remaining alive when the x-spin is up, then at the conclusion of the measurement the state of the cat and the electron will not be one to which the Dieks-Healey account assigns a value to the cat's aliveness.)

It should be clear that non-accurate measurements also create a problem for van Frassen's account. The Hamiltonians which describe non-accurate measurements will typically not satisfy van Frassen's conditions for B's measuring A for any B and A. It will follow that no observables (in addition to ones defined in accordance with the standard prescription) will be well defined. Or if it should turn out that there is a pair of observables which satisfy the condition they will almost always be the wrong ones.

There is a response to the problem created by non-accurate measurements which is suggested by some of Dieks' remarks. What he seems to think is that while the position of the pointer is not well-defined in BCAT*, nonetheless, if the probability of the measuring device making an error is small, then the observable which is well-defined, call it Q*, will be so near the pointer's position that we may consider the pointer as having a position "for all practical purposes." The distance or closeness between two observables can be measured by the degree to which they commute.

But there are at least four things seriously wrong with this bit of FAP reasoning:

1) We want Schroedinger's cat (as we said before) to be either dead or alive! It isn't at all clear what good it might do (that is, it isn't clear what practical purposes it can possible serve) for the cat to be in an eigenstate of some observable that is close to the aliveness observable.

2) In any event (even if there are purposes that would be served by the cat's being in an eigenstate of some observable that is close to the aliveness observable) it isn't even true that if the probability of the measuring device making an error is small then the observable that is well-defined will always be one that is close to

the position of the pointer (or, analogously, the aliveness of the cat). There are Hamiltonians describing almost perfectly accurate measurements which result in final states in which the observable that is well-defined is almost maximally non-commuting with the observable corresponding to the pointer's position (for example, in a state like CAT* in which c_1 is almost equal to c_2).

3) Moreover, what we expect of a solution to the measurement problem is that pointers always end up in determinate positions no matter how inaccurate those measuring devices may be. And R1 certainly does not entail that!

4) And of course there are interactions in which a cat may find itself which are not measurements at all and in which Schroedinger evolution gets the cat into states to which R1 fails to assign a definite value to the aliveness observable or to any observable which is close to it.

We conclude that, contrary to initial appearances, the modal interpretation (and variants of it) fails to guarantee that measurements always have outcomes-indeed, if it were true, then the kinds of measurements that are actually made would almost never have outcomes. Schroedinger's cat has not yet been delivered either dead or alive.

2. The GRW Theory

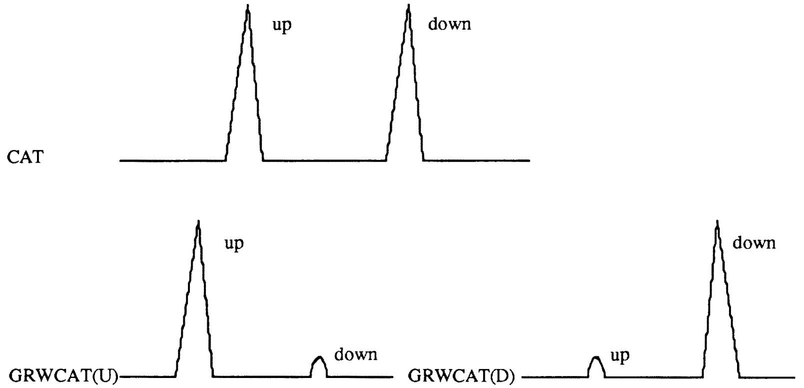
The second proposal we want to examine is an objective theory of the collapse of the wave function due to Ghirardi, Rimini, and Weber which we will call "GRW" (Ghirardi, Rimini, and Weber 1986). It has recently been refined and advocated by John Bell (Bell 1987). Its basic idea is that the Schroedinger equation is not exactly correct. The wave function of an N-particle system $\$(r_1,r_2,...,t)$ usually evolves in accordance with Schroedinger's equation; but every so often (once in $1/N x 10^{15}$ seconds), at random, but with a fixed probability per unit time, the wave function is multiplied by a Gaussian (and normalized). The form of the Gaussian is

G) $K \exp(-[x-r_k]^2/2D^2$

where r_k is chosen at random from the arguments and the width of the Gaussian D is of the order of 10^{-5} cm. The probability of this Gaussian's being centered at any particular point x is stipulated to be proportional to the absolute square of the inner product of $\$(x,...,t)$ (evaluated at the instant just prior to the jump) with G. The effect of the multiplication is a "collapse" of the wave function to a more localized wave function. The probability of a jump per particle (once in 10^{15} seconds) and the width of the multiplying Gaussian (10^{-5} cm) are taken to be new fundamental constants.

What immediately strikes one about the GRW theory (and what is so nice about it) is that unlike von Neuman's collapse theory, it proposes a uniform dynamical law for all interactions and one in which the notion of "measurement" does not appear. Although the theory has in principle empirical differences with any non-collapse view (unlike the modal interpretation previously considered) it is so designed that these differences will be very difficult to detect and have certainly not been detected up until now. For isolated microscopic systems (i.e. systems containing a small number of particles) `jumps' will be so infrequent as to be unobservable in practice; and D (the width of the Gaussian) has been chosen sufficiently large so that violations of the conservation of energy (which jumps produce since they localize wave functions) will be sufficiently small as to be unobservable (over reasonable time intervals) even in macroscopic systems.

Of course the crucial question for the measurement problem is whether GRW collapses produce states in which observables like the pointer's position, the cat's aliveness, and so forth are well defined. Here is GRW's story about that. Consider, for example, the spin measurement previously discussed in which the x-spin of an electron is recorded by the position of a pointer. At the conclusion of the measurement, the state of the pointer is CAT. But on GRW CAT is very unstable. In a very short time (about $1/N10^{15}$ seconds where N is the number of particles which for our measuring device will be on the order of 10^{20})the wave function for the pointer + electron will get multiplied by a GRW Gaussian and a GRW jump will occur and one of the terms in CAT will all but disappear. We can picture the effect of a collapse as follows:



After the jump M+e will be either in GRWCAT(U) or GRWCAT(D) in which almost all the amplitude of the pointer+electron state will be concentrated at either the U or the D position. This seems, at least initially (we will question it later), to justify the claim that after the jump the pointer possess a definite position.

Notice how the fact that the measurement apparatus is macroscopic enters this account. Since the macroscopic pointer contains many particles the probability that there will be a jump to either GRW(D) or GRW(U) in a small time is very great, whereas for microscopic systems the probability of a jump is very small. Notice also that, unlike the standard view, there is a smooth transition from the microscopic to the macroscopic and the laws predict the interactions for measurement devices rather than assume them. And unlike the modal interpretation no special difficulty is caused by measurements which are not perfectly accurate. GRW predicts that the state CAT* will in a short amount of time collapse to a state in which almost all the amplitude is concentrated in one of its four components. If the collapse is to either $|\uparrow>|U>$ or $|\downarrow>|D>$ the measurement is an accurate one. Otherwise it is inaccurate and subsequent measurements may reveal the error. And, of course, Schroedinger's cat, being a macroscopic cat, will occupy the no cat's land between life and death for only an instant.

But there are some difficulties with the GRW theory. One of them is that this theory supposes that anything which deserves to be called a "measurement" involves the correlation of the measured observable with the *position* of some macroscopic object. The trouble is that there are certain very ordinary sorts of measurements for which this just isn't so, and (consequently) the GRW theory won't produce determinate outcomes for these sorts of measurements. Since this difficulty has been developed elsewhere (Albert and Vaidman 1989) we will not further discuss it here. We do, however, want to discuss another problem.

Our worry is that GRW collapses almost never produce definite outcomes even when outcomes are recorded in distinct positions of macroscopically many particles. The reason is that a GRW jump does not literally produce a collapse into an eigenstate of position. It does not collapse a state like CAT into one of its component states. A GRW collapse of CAT yields one of the states like GRWCAT(D) with tails in which *almost* all the amplitude is concentrated in the region around one of the two components but there is non-zero, though very small, amplitude associated with other regions. GRW collapses do not, and cannot, make the other component of the wave function vanish entirely. A GRW collapse is represented mathematically by multiplication of CAT by a Gaussian which is centered at one of the possible positions of the elementary particles in CAT. Since this Gaussian, although very narrow (recall its width is 10^{-5} cm.) has non-zero value everywhere and although it produces a concentration of amplitude around its mean, the resulting state still has non-zero amplitude at every point at which it previously had non-zero amplitude. And it is very important to the theory that the multiplying Gaussian has non-zero amplitude everywhere. If it didn't then the theory would make predictions which we know are contrary to those of quantum mechanics (in particular, it would lead to conflicts with conservation of energy). This means that the post collapse state is not an eigenstate of position (but of some observable which is close to position) and so does not actually assign a definite position to the pointer.

The trouble with the GRW theory, then, is similar to problem number (1) encountered by the modal interpretation. In ordinary measurements the observable that ends up being well-defined on GRW (as on the modal interpretation) is not quite the one we want. (Though GRW ha no special problem with inaccurate measurements.) But problems (2), (3) and (4) for the modal interpretation are not problems for the GRW theory. It (unlike the modal interpretation) genuinely almost gets things right. What GRW is in need of is some way of saying precisely what sort of purpose it serves to almost get things almost right in the way it does. What GRW needs is an argument which shows that it suffices to solve the measurement problem to get pointers almost into eigenstates of position and cats almost into eigenstates of aliveness. Alas, in so far as we are able to judge, that argument may prove very difficult to cook up.

Notes

[1]We would like to thank Hartry Field for the title and Yakir Aharanov for discussions of the modal interpretation.

References

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

_______________. (1989), "Two No-Collapse Interpretations of Quantum Mechanics", *Nous*, 12: 121-38.

Albert, D. and Vaidman, L. (1988), "On a Theory of the Collapse of the Wave-Function", Proceedings of George Mason conference on the foundations of quantum mechanics, to appear.

Bell, J. (1987), *The Speakable and the Unspeakable in Quantum Theory*. Cambridge.

_ _ _ _. (1989), Talk given at a conference "Thirty Years of Uncertainty" Erice, Sicily, August 1989.

Bohm, D. (1952), "A Suggested Interpretation of Quantum Theory in Terms of 'Hidden Variables': Part I," *Physical Review* 85: 180-93.

Dieks, D. (1989), "Resolution of the Measurement Problem Through Decoherence of the Quantum State", *Physics Letters* A 142: 439-46.

Ghirardi, G., Rimini, A., and Weber, M., (1986), *Physical Review* D 34: 470-9.

Healey, R. (1988), *The Philosophy of Quantum Mechanics*. Cambridge.

Kochen, S. (1985), "A New Interpretation of Quantum Mechanics," in *Symposium on the Foundations of Modern Physics*, Lahti and Middlestadt (eds.). Teaneck, NJ: World Scientific Publishing Co., pp. 151-170.

Van Fraassen, B. (1981), "A Modal Interpretation of Quantum Mechanics," in *Current Issues in Quantum Logic*, Beltrametti, E. and Van Frassen, B. (eds.), New York: Plenum, pp. 229-58.

_ _ _ _ _ _ _ _ _ _. (1990), *Quantum Mechanics: an Empiricist View*, ms.

Von Neuman, J.(1955), *The Mathematical Foundations of Quantum Mechanics*. Princeton.