International Journal of Theoretical Physics, Vol. 24, No. 1, 1985

Quantum Theory as a Universal Physical Theory

David Deutsch[1]

Received March 22, 1984

The problem of setting up quantum theory as a universal physical theory is investigated. It is shown that the existing formalism, in either the conventional or the Everett interpretation, must be supplemented by an additional structure, the "interpretation basis." This is a preferred ordered orthonormal basis in the space of states. Quantum measurement theory is developed as a tool for determining the interpretation basis. The augmented quantum theory is discussed.

1. INTRODUCTION

The conventional view of quantum theory is that its logical structure restricts in a fundamental way its domain of potential applicability. It is believed that a quantum theory has no predictive power unless the quantum system under consideration is in interaction with an external system to which quantum theory does not apply (variously called a "macroscopic observable," the "classical level," the "observer," or an "irreversible process"). In that case, quantum theory would be meaningless or inconsistent if regarded as a universal physical theory. (Whether or not a theory is universal is a purely logical property of the theory. This must be distinguished from the empirical question whether quantum theory is universally true. The latter question is not addressed in this paper.) The purpose of this paper is to attack this view by showing how quantum theory may be formulated in a completely intrinsic way; that is, in such a way that every entity mentioned in the formulation has a realization within the quantum formalism. In particular the notion of "measurement," which is required for testing the theory, has such a realization.

A reformulation of quantum theory allowing its interpretation as a universal theory is highly desirable for practical, philosophical, and heuristic

[1]Center for Theoretical Physics, The University of Texas at Austin, Austin, Texas 78712 and Oxford University, Department of Astrophysics, Oxford, England.

reasons. In quantum gravity and quantum cosmology, where the quantum system under consideration is necessarily the whole universe, the conventional interpretation, which applies only to subsystems, fails. Were the technical difficulties in these branches of physics to be overcome today, it would literally not be known what the expressions in the working theory denoted. Epistemologically, it is always desirable to increase the domain of applicability of a theory, in order to make it more testable. Moreover, their attempts to see in the very inadequacy of the conventional interpretation of quantum theory a deep physical principle have often led physicists to adopt obscurantist, mystical, positivist, psychical, and other irrational world views. Undermining, as it thus does, the view that it is the task of physics to seek a systematic understanding of a real, objectively existing world, the widespread acceptance of the conventional interpretation cannot but have impeded the growth of knowledge in physics. As a small illustration, the conjectures made in the last section of this paper, be they true or false, could not be expressed in a theory that is not both realistic and universal.

The major step toward a universal quantum theory was taken in 1957 by Everett (1957) with his "many-universes" interpretation. This is described in Section 5, together with some minor improvements. Everett's interpretation has the advantage over the conventional one (Section 4) that it is realistic and does not require the existence of anything like a "classical level." Unfortunately, and notwithstanding Everett's claim to the contrary, both interpretations share a defect which makes them fall short of universal applicability: They both require *a priori* metaphysical rules for the interpretation of measurements. These rules specify a preferred basis, which I call the *interpretation basis* in the space of physical states. Given a complete specification of a quantum cosmology, for example, but no interpretation basis, neither interpretation could assign a meaning to the abstract theory. The remedy which I propose in Section 7 is that an extra rule be appended to the quantum formalism, specifying how the interpretation basis depends (solely) at each instant on the physical state and dynamical evolution of a quantum system. The *lacuna* in the Everett interpretation would thereby be filled, and it would then be truly universal. The same would be true of the conventional interpretation if it did not, in addition, have other *lacunae* which I believe it is hopeless to try to fill.

My approach toward a general intrinsic specification of the interpretation basis is heuristic, based on an idea of Everett (private communication) that at any rate *during measurements* the basis is determined by the requirement that in that basis the interaction indeed take the form of a measurement. Thus, prior to the discussion of rival interpretations and the interpretation basis, I develop in Sections 2 and 3 a theory of quantum measurement processes along the lines of DeWitt (1968, 1973), and this is used as a heuristic tool throughout the rest of the paper. Section 2 contains a brief statement of the formalism of quantum theory.

In Section 8 I describe a thought experiment whose main purpose is to show how the conventional and Everett interpretations are in principle experimentally distinguishable.

In Section 9 I show how the quantum formalism, if enhanced by the incorporation of an interpretation basis rule, has, quite apart from its universality, a richer structure than before and can be applied to physical problems previously beyond its scope.

2. QUANTUM THEORY

A physical theory consists of two parts, an abstract *formalism* and a concrete *interpretation.* The formalism of quantum theory may be generated by a set of axioms such as the following (d'Espagnat, 1976):

Axiom 1. The world is faithfully represented by a unit vector $|\psi\rangle$ (the *state vector,* or *state*) in a Hilbert space $\mathscr{H}$ (the *state space*), together with a functional $\hat{S}$ (the *action functional*) of Hermitian operators $\hat{\phi}$ acting on H.

Axiom 2. The operators $\hat{\phi}$ are indexed by a set of parameters, including at least one continuous parameter.

Axiom 3. The dependence of $\hat{\phi}_i$ upon the parameters i is determined by dynamical equations obtained from the following variational principle:

$$\frac{\delta \hat{S}[\hat{\phi}_i]}{\delta \phi_j} = 0 \tag{1}$$

where $\delta/\delta\phi_j$ denotes functional differentiation with respect to c-number variations in the form of $\hat{\phi}_j$:

$$\hat{\phi}_j \to \hat{\phi}_j + \delta\phi_j \tag{2}$$

The notation introduced in Axioms 1, 2, and 3 is of course suggestive of quantum field theory, with the parameters i standing for space-time coordinates x^μ, tensor indices and internal symmetry indices, but it is applicable to any quantum theory. For all known quantum theories, the action $\hat{S}$ is in fact a c number, being identical with the action for some classical theory. But one of the objects of this paper is to show that both the formalism and the interpretation of quantum theory can be set up without reference to classical physics, so it is natural to allow for the generality of (1). Boundary conditions on $\hat{\phi}_i$ might be appended to (1) for convenience, or alternatively they can be generated by additional terms in $\hat{S}$.

I shall assume that the action functional is such that the dynamical equations (1) may be integrated to give a first-order system generating unitary time evolution:

$$\frac{\partial \hat{A}_i(t)}{\partial t} = i[\hat{H}(t), \hat{A}_i(t)] \tag{3}$$

The "time" t is the continuous parameter referred to in Axiom 2. $\hat{A}_i(t)$ is any operator at time t. $\hat{H}(t)$ is the Hamiltonian operator.

Measurement theory is usually discussed in the Schrödinger or the interaction picture, but following DeWitt, I have chosen to use the Heisenberg picture because it is more natural both to relativity and to field theory.

The axioms 1, 2, and 3 are far from complete. They give only a framework for the construction of quantum theories. Each choice of the action functional and, if you like, the state vector, gives rise to a particular instance of a quantum theory. However, the formalism generated by Axioms 1, 2, and 3 is already more than sufficient for the construction of a quantum theory of measurement.

Nor are the axioms rigorous. A real axiomatization of quantum theory is far beyond the scope of this paper. My "axioms" would perhaps be better called "mnemonics."

When the quantum system has an infinite number of degrees of freedom (i.e., in field theory) the operators $\hat{\phi}_i$ may have to be distributions (see Gel'fand and Shilov, 1968), the Hilbert space $\mathscr{H}$ may have to be "rigged" (Bohm, 1980), the action may have to be renormalized, and so forth. We shall avoid all these issues by assuming not merely that the world has a finite number of degrees of freedom, but that its state space is finite dimensional. The reason why we can ignore, or rather take for granted, the changes in the formalism necessitated by the passage to an infinite number of degrees of freedom is that these changes are always chosen precisely so as to *restore*, not destroy, the relevant properties of a finite-dimensional theory. I must admit immediately that I have no proof that any of my results or conclusions possess infinite-dimensional analogs, though I conjecture that they all do.

It is worth pointing out that the term *state* in Axiom 1 is used exclusively to denote a "pure" state of the whole world, and not a "mixed state" of some ensemble or of some part of the world. This is in line with the general philosophy of this paper, but not with the usage of many authors. In referring to ensembles I shall follow d'Espagnat (1976) and use the word *case* ("pure case," "mixed case") instead of *state*.

The interpretation of a theory provides a mapping between the formalism and objective reality. It asserts for every element of some universe of discourse (the "world") that it is denoted by a particular element of the formalism. For the theory to be a physical theory the universe of discourse must contain (though *not* exclusively) some elements of ordinary experience so that the theory may be testable. The wider and the more naive is the notion of "ordinary experience" used by the interpretation, the more general and the deeper is the theory.

Again, for the limited purposes of quantum measurement theory a very incomplete set of interpretational axioms suffices. These can be chosen so as to be to a surprising extent independent of the controversy surrounding the interpretation of quantum theory:

Axiom 4. The world may be divided into *subsystems* which have state spaces of their own. $\mathscr{H}$ is the direct product $\mathscr{H}_1 \otimes \mathscr{H}_2 \otimes \cdots$ of the state spaces of the subsystems.

Axiom 5. Hermitian operators on $\mathscr{H}$ correspond to *observables.*

Axiom 6. When $|\psi\rangle$ is an eigenvector of an observable $\hat{O}$ with eigenvalue $\lambda = \langle\psi|\hat{O}|\psi\rangle$ then $\hat{O}$ *possesses* the value λ. If the observable were to be (perfectly) measured then the result of the measurement would invariably be λ.

Axiom 7. Observables not assigned values by Axiom 6 do not possess definite values.

Axiom 4 has a similar function to Everett's (DeWitt and Graham, 1973) "postulate of complexity," that "the world is decomposable into systems and apparata." This is obviously necessary for measurement theory, but the axiom also has a wider role, in allowing quantum theory to be applicable to certain *parts* of the world instead of just to the whole. We shall however see in Section 7 that this direct product construction is not the only way in which the world may be divided into subsystems.

Axiom 6 is of great practical importance in measurement theory because it is uncontroversial (i.e., the rival interpretations of quantum theory agree on it). Unfortunately, it is not self-contained since it refers to the "result" of a potential measurement, a concept about which, we shall see, there is controversy. For the quantum theory defined here to be a universal theory, either this concept ("the result") will have to be given meaning within the formalism, or Axiom 6 must be replaced by a more general axiom of which it is a limiting case. In the Everett interpretation (Section 6) the latter alternative is taken.

An "observable" (Axiom 5) is something which could be measured by a measurement, if a suitable apparatus were present at the right place(s) and time(s). In this paper I shall not discuss the interesting question *which* thus informally defined observable corresponds to *which* operator. What a "measurement" and an "apparatus" are will, I hope, emerge in the following sections.

3. SUBSYSTEMS OF QUANTUM SYSTEMS

Axiom 4 establishes that quantum systems may be described in terms of their subsystems. This is convenient from the point of view of testing quantum theory experimentally, since the very concept of "measurement" requires that at least a "system" and an "observer" exist, and in a universal physical theory these must both be subsystems of the world. It is convenient to summarize here the formalism and terminology which I shall be using to describe subsystems. To this end, let us divide the world into two subsystems. Then the state space $\mathscr{H}$ of the world is identified with the $n_1 n_2$-dimensional direct product $\mathscr{H}_1 \otimes H_2$ of the subsystems' state spaces. In order to preserve the vector space structure, this identification must take the form of a linear mapping

$$\mathbf{L} : \mathscr{H} \to \mathscr{H}_1 \otimes \mathscr{H}_2 \tag{4}$$

That is, if

$$\begin{aligned} &|a_1\rangle \quad (1 \le a_1 \le n_1) \\ &|a_2\rangle \quad (1 \le a_2 \le n_2) \\ &|\alpha\rangle \quad (1 \le \alpha \le n_1 n_2) \end{aligned} \tag{5}$$

are arbitrary orthonormal bases in $\mathscr{H}_1$, $\mathscr{H}_2$, and $\mathscr{H}$ respectively, then $\mathbf{L}$ is represented by a bivector $\mathscr{L}^{\alpha}_{a_1 a_2}$ where

$$\mathscr{H}_1 \otimes H_2 \ni |a_1\rangle\, |a_2\rangle \leftrightarrow \sum_{\alpha} \mathscr{L}^{\alpha}_{a_1 a_2} |\alpha\rangle \in \mathscr{H} \tag{6}$$

and is unitary in the sense

$$\begin{aligned} &\sum_{\alpha} \mathscr{L}^{\alpha}_{a_1 a_2} L^{t_1 t_2}_{\alpha} = \delta^{t_1}_{a_1} \delta^{t_2}_{a_2} \\ &\sum_{a_1 a_2} \mathscr{L}^{\alpha}_{a_1 a_2} \mathscr{L}^{a_1 a_2}_{\beta} = \delta^{\alpha}_{\beta} \end{aligned} \tag{7}$$

Here and throughout this paper, raising and lowering of state space indices denotes complex conjugation

$$X_{\alpha} \equiv (X^{\alpha})^* \tag{8}$$

Two such mappings $L^{\alpha}_{a_1 a_2}$ and $M^{\alpha}_{a_1 a_2}$ are said to define the same *product structure* on $\mathscr{H}$ (i.e. they decompose $\mathscr{H}$ into the same two subspaces $\mathscr{H}_1$ and $\mathscr{H}_2$) whenever they are related by unitary transformations $U^{b_1}_{a_1}$ and $V^{b_2}_{a_2}$

confined to $\mathscr{H}_1$ and $\mathscr{H}_2$, i.e. when

$$\mathscr{L}^{\alpha}_{a_1 a_2} = \sum_{b_1 b_2} U^{b_1}_{a_1} V^{b_2}_{a_2} M^{\alpha}_{b_1 b_2} \tag{9}$$

An observable $\hat{X}$ on $\mathscr{H}$ is said to "be confined to system 1," or to "be a system 1 observable" if

$$\sum_{\alpha\beta} \mathscr{L}^{\alpha}_{a_1 a_2} \mathscr{L}^{b_1 b_2}_{\beta} \langle \beta | \hat{X} | \alpha \rangle = X^{b_1}_{a_1} \delta^{b_2}_{a_2} \tag{10}$$

for some $X^{b_1}_{a_1}$. (10) shows that $\hat{X}$ then has a certain degeneracy structure (n_1 sets of n_2 identical eigenvalues). Conversely, if any observable $\hat{X}$ has those degeneracies then they determine a product structure with respect to which $\hat{X}$ is confined to a subsystem.

A classical system is said to be "isolated" whenever there are no external forces acting on it. In quantum theory there are several different notions of "isolation" or "independence" of subsystems. In this paper I shall be using two of them, *dynamical independence*, which is somewhat analogous to "isolation" in classical physics, and *kinematical independence*, which has no classical analog. (All classical systems are kinematically independent.)

If the state of the world $|\psi\rangle$ is simultaneously an eigenstate of some nondegenerate subsystem 1 observable and some nondegenerate subsystem 2 observable then I shall call the subsystems 1 and 2 kinematically independent. In that case, there exist elements $|\xi_1\rangle$ and $|\xi_2\rangle$ of $\mathscr{H}_1$ and $\mathscr{H}_2$ such that $|\psi\rangle \in \mathscr{H}$ can be identified with the product

$$|\xi_1\rangle |\xi_2\rangle \in \mathscr{H}_1 \otimes \mathscr{H}_2 \tag{11}$$

Consequently the subsystems 1 and 2 have all the properties of worlds with states $|\xi_1\rangle$ and $|\xi_2\rangle$. For example, if measurements are made on subsystem 1 by a third kinematically independent subsystem, 3, then the probability distribution function for the results is the same as if subsystems 1 and 3 constituted the whole world. By the same token, if subsystem 1 itself consists of two kinematically independent subsystems and one measures the other, then the probability of any given result is the same as if subsystem 1 were the whole world. Furthermore the results of separate measurements on subsystems 1 and 2, whether external (i.e., made by a third subsystem) or internal, are *uncorrelated*—i.e., their joint probability distribution function is just the product of the distribution functions for the individual measurements.

All these examples referring to measurements and potential measurements may be substantiated from the theory of measurement to be developed in the next section.

I have been careful in the above to avoid making a statistical (ensemble) interpretation of probability statements. It is perfectly legitimate to regard the *probabilities* of the several results of a potential measurement on a system as objective physical properties of that system. The appropriate technical device is to use an interpretation of the abstract calculus of probabilities, namely, a "propensity" interpretation (Popper, 1967)[2] different from the more usual "frequency" or "ensemble" interpretations. Propensity interpretations assign objective meanings to probabilities of events in single systems (which, after all, the world is) instead of in ensembles (which the world is not). This can be done in either the conventional or the Everett interpretation of quantum theory.

Because of the above-mentioned properties of joint measurements on kinematically independent subsystems, the usual term for what I have called "kinematically independent" is "uncorrelated." However, I should like to warn the reader that this meaning of the term "uncorrelated" is sometimes quite different from its meaning in ordinary language. For example, just after a perfect measurement [see equation (17) below], in the case where the state $|\psi\rangle$ happens to be an eigenstate of the observable being measured, the system and the apparatus are kinematically independent, though in ordinary language we should probably call them perfectly *correlated*, certainly not uncorrelated, since their properties are in perfect agreement. For this reason, I shall use the term *uncorrelated* only in cases when the technical meaning agrees with ordinary usage. The same perfect measurement example shows that subsystems can be strongly interacting but remain kinematically independent.

Throughout the extended general discussion of quantum measurement theory, to which I shall continually return in this paper, it may be helpful to have in mind a specific example. The best-known laboratory example of a quantum system with a finite-dimensional state space is a spin-$\frac{1}{2}$ system, such as a silver atom with total angular momentum $\frac{1}{2}\hbar$. The archetypal measurement of this system is the experiment of Stern and Gerlach, an excellent discussion of which is given in Feynman's *Lectures on Physics* (Feynman, 1965). In such an experiment the component $\mathbf{n} \cdot \hat{s}$ of the atom's spin $\hat{s}$ in some desired direction $\mathbf{n}$ is measured by the angle $\hat{\theta}$ at which the atom emerges from an inhomogeneous magnetic field. In reality $\hat{\theta}$ is a continuous observable and has an infinity of eigenvalues, but ideally only two of these values are ever taken by $\hat{\theta}$ after a measurement. They correspond to the eigenvalues $\pm\frac{1}{2}\hbar$ of $\mathbf{n} \cdot \hat{s}$. Thus the relevant eigenstates of $\hat{\theta}$ span a two-dimensional state space. An alternative example is the measurement of $\mathbf{n} \cdot \hat{s}$ by the spin of another atom.

We now resume discussion of a general quantum system. A generic state $|\psi\rangle$ of $\mathcal{H} \leftrightarrow \mathcal{H}_1 \times \mathcal{H}_2$ does not represent kinematically independent

[2] Popper's remarks on quantum theory *per se* are in error.

subsystems. Its expression in terms of any bases $\{|a_1\rangle\}$ and $\{|a_2\rangle\}$ in $\mathscr{H}_1$ and $\mathscr{H}_2$ is not a product $|\xi_1\rangle|\xi_2\rangle$ but a linear superposition of such products

$$|\psi\rangle \leftrightarrow \sum_{a_1 a_2} c^{a_1 a_2}|a_1\rangle|a_2\rangle \tag{12}$$

for some $c^{a_1 a_2}$ satisfying

$$\sum_{a_1 a_2} |c_{a_1 a_2}|^2 = 1 \tag{13}$$

Therefore generically, subsystems of quantum systems cannot be described by states restricted to the subsystems' state spaces. Joint measurements in general show correlations between the subsystems. Indeed they show more: *nonseparability* (d'Espagnat, 1976). That is, the probability distribution function for measurements on subsystem 1 depends not only on the result of the measurement of subsystem 2 (that would just be correlation) but also on *what measurement* is performed on subsystem 2 (something which is "freely specifiable" by the observer). This is true even when the subsystems are not interacting. This phenomenon, which underlies the famous thought experiment of Einstein, Podolski, and Rosen (1935) and Bell's (1964) theorem, is uniquely characteristic of quantum theory. I shall return to it in Section 7, where we shall see why it cannot be used for signaling between noninteracting subsystems.

We have seen that if subsystems are kinematically independent at some instant then at that instant, even if they are interacting, they can be given autonomous descriptions, whereas otherwise they have joint properties which cannot be inferred from their individual properties. The dynamical evolution of the world may be such as to preserve this autonomy of subsystems, or it may not. A sufficient condition is that the Hamiltonian operator $\hat{H}$ [equation 3)] be a sum of operators $\hat{H}_1 + \hat{H}_2$, confined to the subsystems. But a weaker condition, which I shall call "dynamical independence," is both necessary and sufficient to ensure that a given kinematically independent state $|\psi\rangle$ remains so:

$$(\hat{H} - (\hat{H}_1 + \hat{H}_2))|\psi\rangle = 0 \qquad \text{(for some } \hat{H}_1 \text{ and } \hat{H}_2) \tag{14}$$

Stated in words, kinematically independent subsystems are also dynamically independent if the state is an eigenstate of the Hamiltonian *modulo* terms confined to the subsystems.

I shall not require a definition of dynamical independence for general (kinematically dependent) subsystems, though in a sense the interpretation basis construction of Section 7 (53) provides one.

4. MEASUREMENT PROCESSES AND THE MEASUREMENT PROBLEM

A quantum measurement, just like a classical measurement, is a process during which the value of one ("apparatus") observable comes to depend systematically upon the value of another ("system") observable. Thus, in the simplest possible model of a quantum measurement, the world consists of two subsystems: The system being measured, with an n_1-dimensional state space $\mathcal{H}_1$, and the apparatus with an n_2-dimensional state space $\mathcal{H}_2$, where $n_2 \geq n_1$. A system observable $\hat{\phi}_1(t')$ is measured by an apparatus observable $\hat{\phi}_2(t'')$. The subscripts 1 and 2 remind us that the observables are confined to their respective subsystems. The appropriate product structure may be determined from $\hat{\phi}_1(t')$ and $\hat{\phi}_2(t'')$. t' and t'' are particular values of t, the absolute time, the only parameter upon which the observables depend in this model.

In a causal world, $\hat{\phi}_2(t'')$ can measure $\hat{\phi}_1(t')$ only if $t'' > t'$. This restriction is of great practical importance, but is *not* imposed by the structure of quantum theory, but rather by the state of the real world. I shall return to this point in Sections 7 and 9.

In the Stern-Gerlach experiment, the system is the atomic spin $\mathbf{n} \cdot \hat{s}$ and the apparatus is the angle $\hat{\theta}$. Performing a real measurement in the laboratory involves introducing a coupling which in general causes the motion of system and apparatus variables to depend on each other. There are at least three types of quantity in which the experimentalist might in principle be interested: (1) The value of $\hat{\phi}_1$ before the measurement began. This is $\hat{\phi}_1(t')$, as assumed in the model. (2) The value of $\hat{\phi}_1$ at the end of the measurement, i.e., $\hat{\phi}_1(t'')$, and (3) The value that $\hat{\phi}_1$ would have had at t'' if the measurement had not been performed. Possibility (1) is the usual one in a physics laboratory, where we wish to ignore changes which we ourselves have introduced. Possibility (2) is appropriate when the actual condition of the system, however caused, is the subject of interest. (An example is a general electron.) Possibility (3) arises when the measurement itself induces spurious changes (such as spin precession in the Stern-Gerlach experiment) in a quantity which it takes time to measure. Now the results of type (1) and type (3) experiments are both hypothetical constructs. At time t'', the significance of the earlier time t' and the value of the unperturbed system observable resides more in the intention of the observer (see Section 12) than in the objective properties of his experiment. Thus we shall find that it is the analysis of measurements of type (2) which sheds the most light on the foundations of quantum theory. However, in order to keep the discussion of measurement theory *per se* as general as possible, the model measurements I discuss will all be nonperturbing, so that (1), (2), and (3) coincide.

Prior to a perfect measurement (i.e., just before the time t'), the system and the apparatus are kinematically independent. In this section it will suffice to consider only the very special cases where $|\psi\rangle$ is an eigenstate of $\hat{\phi}_1(t')$. Let us assume also, for the sake of simplicity of notation (the assumption is not otherwise necessary) that $|\psi\rangle$ is an eigenstate of $\hat{\phi}_2(t')$. Thus

$$|\psi\rangle = |a_1, t'; \bar{a}_2 t'\rangle \tag{15}$$

where

$$\begin{aligned} [\hat{\phi}_1(t') - \phi_{a_1}]|a_1, t'; a_2, t'\rangle = 0 \\ [\hat{\phi}_2(t') - \phi_{a_2}]|a_1, t'; a_2, t'\rangle = 0 \end{aligned} \tag{16}$$

The ϕ_{a_1} $(1 \le a_1 \le n_1)$ are the eigenvalues of $\hat{\phi}_1(t)$ and the ϕ_{a_2} $(1 \le a_2 \le n_2)$ are the eigenvalues of $\hat{\phi}_2(t)$, with t either t' or t'' since we shall contrive matters so that the eigenvalues of $\hat{\phi}_1$ and $\hat{\phi}_2$ will not change during the measurement. However, the eigen*states* of $\hat{\phi}_1(t)$ and $\hat{\phi}_2(t)$ do change with time, as the method of labeling in (5) reminds us. The particular value $\bar{a}_2$ of a_2 corresponds to the *receptive* value $\phi_{\bar{a}_2}$ of $\hat{\phi}_2(t')$. Loosely speaking, this is the value to which $\hat{\phi}_2(t')$ must be "set" in order to switch on the apparatus. Axioms 4, 5, and 6, incomplete though they are, allow us to interpret (15): "At the time t', $\hat{\phi}_1$ possesses the value ϕ_{a_1} and $\hat{\phi}_2$ possesses its receptive value $\phi_{\bar{a}_2}$."

During the measurement (i.e., at times between t' and t'') the dynamical evolution of $\hat{\phi}_2$ is such that it comes to possess a value which depends on that of $\hat{\phi}_1(t')$. Specifically, if

$$|\psi\rangle = |a_1, t''; A_2(a_1), t''\rangle \tag{17}$$

where $A_2(a_1)$ is an assignment of a distinct value of a_2 to each value of a_1, then a perfect measurement has taken place. This follows from the (Axiom 6) interpretation of (4.7): "At the time t'', $\hat{\phi}_1$ still possesses the value ϕ_{a_1}, but $\hat{\phi}_2$ possesses the value $\phi_{A_2(a_1)}$." Moreover, the possession of this value by $\hat{\phi}_2(t'')$ indicates uniquely which value $\hat{\phi}_1$ possesses. The function A_2 may be thought of as a "calibration" of the apparatus.

The measurement described by equations (4.1) and (4.7) is "perfect" in two senses. Firstly, it is *accurate*, in the sense that each initial value of the system observable determines a different final value of the apparatus observable. Secondly, as promised above, it is *nonperturbing*, i.e., the system observable is unaffected by the measurement interaction. These properties illustrate two interesting implications of quantum measurement theory: Firstly that quantum theory gives rise to no absolute restriction upon the accuracy with which a single observable can be measured (Bohr and Rosenfeld, 1933; DeWitt, 1933, 1968; DeWitt and Graham, 1933). Secondly that a quantum system can act without itself being acted upon. Since these implications are slightly counterintuitive, and as a preparation for the following section, I shall now show that an interaction (i.e., an action) exists which would generate the dynamical evolution constituting the perfect measurement (3.1), (3.3).

In view of (2.3), observables at different times are related by a unitary transformation

$$\hat{O}(t'') = e^{-i\hat{X}}\,\hat{O}(t')\,e^{i\hat{X}} \tag{18}$$

Therefore the eigenstates of $\hat{\phi}_1$ and $\hat{\phi}_2$ at different times are related by

$$|a_1, t''; a_2, t''\rangle = e^{-i\hat{X}}|a_1, t'; a_2, t'\rangle \tag{19}$$

Regarded as conditions on the matrix elements of $e^{i\hat{X}}$, (15), (17), and (19) amount to

$$\langle b_1, t'; b_2, t'|e^{i\hat{X}}|a_1, t'; \bar{a}_2, t'\rangle = \delta^{b_1}_{a_1}\delta^{b_2}_{A_2(a_1)} \tag{20}$$

Let $A_2(a_1, a_2)$ be any function which for fixed a_1 is a permutation of the integers 1 to n_2, such that $A_2(a_1, \bar{a}_2) = A_2(a_1)$. Then it is easy to show that

$$(b_1, t'; b_2, t'|e^{i\hat{X}}|a_1, t'; a_2, t'\rangle = \delta^{b_1}_{a_1}\delta^{b_2}_{A_2}(a_1, a_2) \tag{21}$$

is a solution of (20), and is unitary. The unitary transformation $e^{i\hat{X}}$ may of course be generated by the Hamiltonian

$$\hat{H} = (t'' - t')^{-1}\hat{X} \tag{22}$$

acting at times between t' and t''. The stationarity of the following quantum action functional with respect to c-number variations $\delta\phi_i(t)$ of $\hat{\phi}_i(t)$

$$\hat{S}[\hat{\phi}_i(t)] = \frac{1}{2}\sum_{i=1}^{2}\int_{t'}^{t''}\left\{\left(\frac{\partial\hat{\phi}_i}{\partial t}\right)^2 + \frac{1}{2}[\hat{H}, \{\hat{\phi}_i, [\hat{H}, \hat{\phi}_i]\}]\right\}dt \tag{23}$$

together with the initial condition

$$\left.\frac{d\hat{\phi}_i}{dt}\right|_{t=t'} = i[\hat{H}, \hat{\phi}_i(t')] \tag{24}$$

reproduces the desired motion (18) of all observables, where $\{\hat{A}, \hat{B}\}$ denotes the anticommutator $\hat{A}\hat{B} + \hat{B}\hat{A}$. Notice in passing that any dynamical motion for an arbitrary quantum system can be generated by a stationary quantum action principle of the form (23), (24). Thus, although it may be convenient, it is never in principle necessary for an associated classical theory to appear in the construction of a quantum theory.

It follows from equations (21) to (24) that the requirement that the measurement interaction be perfectly accurate and nonperturbing is not inconsistent with quantum theory. This is not to say that couplings such as (23) are necessarily available in the laboratory! Nevertheless, (23) is as far as any complete discussion of quantum measurement theory need go until the nature of the quantum action functional of the real world is understood at a less phenomenological level than it is at present.

The model described in this section is idealized in another sense also. A real laboratory apparatus does not consist of just one observable like $\hat{\phi}_2(t'')$, but rather a long chain of them, each measuring the previous one and each more "macroscopic" than the previous one, ending with the brain of the observer. This chain of measurements is sometimes called a "complete" measurement and its links "elementary" measurements.

4.1. The Measurement Problem

In the preceding discussion of a restricted class of measurement processes, where the state $|\psi\rangle$ of the world is an eigenstate of the observable $\hat{\phi}_1(t')$ being measured, we found that the requirement that the measurement be accurate and nonperturbing essentially determines the dynamical evolution law [equations (18) and (20)]. But this law determines equally well what happens in a more general case when $|\psi\rangle$ is not an eigenstate of $\hat{\phi}_1(t')$. Let us suppose that the model apparatus described above has been prepared and switched on properly, i.e., that the system and apparatus are uncorrelated at time t' and that the apparatus is receptive, but that $|\psi\rangle$ now has no special property with respect to $\hat{\phi}_1(t')$. The most general state with these properties is

$$|\psi\rangle = \sum_{a_1} c^{a_1} |a_1, t'; \bar{a}_2, t'\rangle \tag{25}$$

with

$$\sum_{a_1} c^{a_1} c_{a_1} = 1 \tag{26}$$

In the Stern–Gerlach experiment, $|\psi\rangle$ might be an eigenstate of some spin component which does not commute with the one being measured. In view of (13), we have

$$|\psi\rangle = \sum_{a_1} c^{a_1} |a_1, t''; A_2(a_1), t''\rangle \tag{27}$$

In general, this is not an eigenstate of the apparatus observable $\hat{\phi}_2(t'')$. Indeed it is not in general an eigenstate of *any* system or apparatus observable at the time t''. Here is the origin of the "problem of measurement" in quantum theory. In the case where all but one of the coefficients c^{a_1} vanished, we could use the "uncontroversial" interpretation Axiom 6 to say what the *result* of measuring $\hat{\phi}_1(t')$ was. It was $\phi_{A_2(a_1)}$. But now the state (27) contains each of the n_1 possible outcomes of the measurement on an equal footing. Moreover, because of the vector space structure of $\mathscr{H}$ and the linearity of the dynamical evolution laws, no construct either in, or consistent with, the formalism of quantum theory is capable of distinguishing between these outcomes. On the one hand, this is to be expected. The apparatus was, after all, designed to measure the value of $\hat{\phi}_1(t')$, which, if the state is (25), simply does not possess a definite value (Axiom 7). On the other hand, the most elementary intuition tells us that if we nevertheless set the apparatus in motion, we shall always observe it to record a definite value. And it is axiomatic (and experimentally corroborated in, say, the Stern-Gerlach experiment) that this will be one of the eigenvalues. *Which one?*

5. THE COPENHAGEN "INTERPRETATION"

The greatest sophism in the history of science (H.D. Zeh, 1980)

One widespread solution to this problem involves a fundamental change in the quantum formalism presented in Section 2. The idea is that the state $|\psi\rangle$, which according to the formalism of Section 2 never changes, is in fact subject to an intermittent, discontinuous motion. At certain instants t_i, at which measurements are said to have been "completed," $|\psi\rangle$ changes into a randomly chosen simultaneous eigenstate of the observable being measured and the observable doing the measuring. The probability of the eigenstate $|a_1, t_i; a_2, t_i\rangle$ being chosen is

$$|\langle\psi_{i-1}|a_1, t_i; a_2, t_i\rangle|^2 \tag{28}$$

where we have denoted by $|\psi_i\rangle$ the state of the world between the ith and $(i+1)$th completion of a measurement. The point of all this is that the "uncontroversial" Axiom 6 can then again be used to interpret the state of affairs just after any of the instants t_i: "The completed measurement always has one actual result, namely the eigenvalue corresponding to the randomly chosen eigenstate."

The version of quantum theory with this stochastic law of motion for $|\psi\rangle$ we call the (objective) "collapse interpretation" ("C.I.", for short), though, as we shall see, the word "interpretation" is very much a misnomer. The C.I. postulates a different *formalism* from Section 2. The interpretation in Section 2 is neither changed nor augmented by the C.I.

No complete formulation of the C.I. has yet been achieved. This is mainly because of the notorious difficulty in finding a criterion for specifying the preferred instants t_i; for specifying, in other words, when a measurement is "completed." In order to demonstrate this difficulty, let us make the hypothesis that the instant t'' at which our model measurement ends is in fact one of the t_i. Then (27) holds just before t'', but at the next instant

$$|\psi(t > t'')\rangle = |\bar{a}_1, t''; A_2(\bar{a}_1), t''\rangle \tag{29}$$

for some particular $\bar{a}_1$ chosen randomly with probability

$$P(\bar{a}_1) = |\langle\psi(t < t'')|\bar{a}_1, t''; A_2(\bar{a}_1), t''\rangle|^2 \tag{30}$$

Now this is to some extent in accord with experiment, for the interpretation the apparatus records the value $\phi_{\bar{a}_1}$ for $\hat{\phi}_1(t')$, and this is indeed what is observed, with relative frequency $P(\bar{a}_1)$. By "observed," though, we can only mean "measured." That is, if the system and apparatus observables at t'' are themselves measured by a third subsystem, the "observer," at time $t''' > t''$, say, then they are observed to have definite values such as $\phi_{\bar{a}_1}$ and $\phi_{A_2(\bar{a}_1)}$, respectively. Thus, the empirical observation has, after all, no bearing on whether or not t'' was one of the t_i: (29) might still be wrong, and $|\psi\rangle$ might have kept its original form up to the time t''', and only then become

$$|\psi(t > t''')\rangle = |\bar{a}_1, t''; A_2(\bar{a}_1), t''; A_3(\bar{a}_1, A_2(\bar{a}_1)), t'''\rangle \tag{31}$$

The $A_3(\bar{a}_1, A_2(\bar{a}_1))$'th eigenvalue of $\hat{\phi}_3(t''')$ is the one which corresponds to the observer's having observed $\hat{\phi}_1(t')$ to possess the value ϕ_{a_1} and $\hat{\phi}_2(t'')$ the value $\phi_{A_2(\bar{a}_1)}$. I have written (31) in a mixed basis of eigenstates of observables at different times [viz. $\hat{\phi}_1(t''$ or $t')$, $\hat{\phi}_2(t'')$, $\hat{\phi}_3(t''')$] in order not to have to make the assumption that the second measurement ("observation") is nonperturbing.

So, the measurement may or may not be complete at t'', but by t''' it must be. Or must it? Our only reason for believing that $\hat{\phi}_1(t'')$, $\hat{\phi}_2(t'')$ and $\hat{\phi}_3(t''')$ possess values after t''' is the evidence of observers like $\hat{\phi}_3(t''')$ who say they get definite results with the right probabilities (5.3). But our acquisition of such evidence at time $t'''' > t'''$ is itself a measurement, essentially of $\hat{\phi}_3(t''')$, and so by a repetition of the above argument we have no reason to believe that any discontinuous change was suffered by $|\psi\rangle$ until the time t''''. Furthermore, our very memory of our own scrutiny of the evidence at t'''' is yet another measurement of our own brain state (or notebook), performed at some still later time. Thus, in our model, suitably extended to include all measurements and reports of measurements of $\hat{\phi}_1(t')$ up to the present instant, discontinuous changes in $|\psi\rangle$ can be postulated for any time, or sequence of times starting with t', at which an elementary measurement ended. Yet, if this were the whole story, the C.I. would be quite viable even without specifying the t_i, because it would not make any physical difference which elementary measurements were designated "complete" and which were not, so long as at least one was complete between t' and now. However, by using a more complicated apparatus, it is possible to detect experimentally whether a given measurement in the past was complete or not (i.e., whether $|\psi\rangle$ changed or not). Experiments of this general type are called *interference* experiments, and I shall now describe one appropriate to our model.

Let the state at time t' be (25) and let the elementary measurement of $\hat{\phi}_1(t')$ by $\hat{\phi}_2(t'')$ be performed as described in Section 4. Let $\hat{\xi}_1(t)$ be a system observable such that

$$(\hat{\xi}_1(t') - \xi)|\psi\rangle = 0 \tag{32}$$

The desired interference experiment is represented by any interaction which, taking place between the times t'' and t''', gives rise to a unitary transformation which is the inverse of the perfect measurement (21). In the Stern–Gerlach experiment, a system of magnetic lenses might cause the two trajectories precisely to join up again (having described the same path length). If the measurement at t'' was *not* complete, this transformation will restore to all the observables in the theory, in particular $\hat{\xi}_1$ and $\hat{\phi}_2$, the values they had at t'. Thus at time t''' the state would still be

$$|\psi\rangle = \sum c_{a_1}|a_1, t'''; \bar{a}_2, t'''\rangle \tag{33}$$

and

$$(\hat{\xi}_1(t''') - \xi)|\psi\rangle = 0 \tag{34}$$

On the other hand, if the measurement *was* complete at t'', and if the state had consequently changed to (29), then with respect to the t''' basis it could be expressed as

$$|\psi(t > t'')\rangle = |\bar{a}_1, t'''; \bar{a}_2, t'''\rangle \tag{35}$$

A measurement of $\hat{\xi}_1(t''')$ could now distinguish between the alternatives (33) and (35). If the state is (33) then $\hat{\xi}_1(t''')$ possesses the value ξ and therefore a measurement would yield this result with certainty. If the state is (34) then values other than ξ will be observed with nonzero probability. If the identical experiment is repeated many times, and if a value other than ξ is ever observed, then the measurement was complete at time t''.

Notice that if there was no change of state at t'', then the initial measurement, its undoing, and the measurement of $\hat{\xi}(t''')$ are together entirely equivalent to a measurement of $\hat{\xi}(t')$. But if there was such a change, then the value of $\hat{\xi}(t')$ is no longer observable at any time after t''. This loss of information [or creation of new information, if you will (Wheeler, 1977)] is a characteristic feature of the C.I. which was not present in our original (Section 2) version of the quantum formalism.

The necessity of repeating identical measurements in order to test probabilistic predictions is generic in quantum theory. For such experiments it is often convenient to use the *ensemble approximation*: A long finite sequence of measurements on identical systems is approximated by an infinite set of measurements on an ensemble. The properties of a quantum ensemble are described by its density operator $\hat{\rho}$ (Davies, 1976; d'Espagnat, 1976). At the moment of completion of a measurement, the density operator for the ensemble which approximates sequences of our model measurement changes discontinuously from

$$\hat{\rho} = \sum_{a_1 b_1} C^{a_1} C_{b_1} |a_1, t''; A_2(a_1), t''\rangle\langle b_1, t''; A_2(b_1), t''| \quad (36)$$

corresponding to the state (27), to

$$\hat{\rho}(t > t'') = \sum_{a_1} |C_{a_1}|^2 |a_1, t''; A_2(a_1), t''\rangle\langle a_1, t''; A_2(a_1), t''| \quad (37)$$

which is a mixed case corresponding to no pure state.

All interference experiments which have so far been performed have given the result that their intermediate elementary measurement [analogous to that of $\hat{\phi}_1(t')$ by $\hat{\phi}_2(t'')$] was *not* complete. On the other hand, since it is the purpose of the C.I. to give a realization of the apparent single valuedness of our (the observers') experience, every chain of elementary measurements leading to an observation by "us" must have at least one complete member.

Many criteria for "completeness" of a measurement have been proposed informally, e.g., that completeness requires an "irreversible act of amplification" (George *et al.*, 1972; Misra *et al.*, 1979), or registration in the mind of a conscious observer (von Neumann, 1930, 1932/1969, 1946, 1955, 1964; Wigner, 1961/1962), or even in the minds of a whole community of people (London and Bauer, 1939). However, no definite criterion has ever emerged from these proposals. At the risk of adding to the confusion, I should like to propose yet another version of the C.I., which avoids this problem: "At *every* elementary measurement there is a small probability that the state will change discontinuously according to (27) and (29)." This probability may, for example, be a universal constant, chosen to be so small that interference is in practice always observed on a microscopic scale, but large enough so that it is overwhelmingly likely that a long chain of elementary measurements will be complete. Like all versions of the C.I., this one is distinctly *ad hoc*, but it does have the advantage of being definite. By the same token, it is potentially testable and may, one day, be refuted experimentally (see Section 8).

The severity of the "completeness-of-measurements" problem, and the obfuscatory effects of extraphysical attempts to solve it, have overshadowed another important deficiency of the C.I.: Neither the product structure dividing the world into system+observer nor the observables $\hat{\phi}_1(t'')$ and $\hat{\phi}_2(t'')$ to whose eigenvectors the state makes its discontinuous changes are specified by the C.I. Because of the grossness of the physical constitution of all known "observers," this lack of specification does not at present limit the usefulness of the C.I. in the laboratory: Knowledge *external* to quantum theory tells us that the subsystem in the white coat is the observer, the subsystem on the bench is the system under observation, and the observable being measured is the one upon which the motion of the needle on the calibrated scale depends. Thus the C.I. version of quantum theory is not a universal theory until these "rules of thumb," no less indispensable to the C.I. than the criterion for completeness-of-a measurement, are generalized and given expression within the quantum formalism. We shall see in the next section that the Everett interpretation, though completely free of the intractable completeness-of-measurement problem, does share with the C.I. the deficiency just described. In Section 7 the missing structures will be supplied for both the C.I. and Everett formalisms.

6. THE EVERETT ("RELATIVE STATE") INTERPRETATION

> No escape seems possible from this relative state formulation if one wants to have a complete mathematical model for the quantum mechanics that is internal to an isolated system. Apart from Everett's concept of relative states, no self-consistent system of ideas is at hand to explain what one shall mean by quantizing a closed system like the universe... (J. A. Wheeler, 1957)

It is possible to leave the formalism of Section 2 alone and solve the problem of measurement solely by means of an *interpretation.* Such an interpretation was first attempted by Everett (1957).

The problem of measurement, we recall, arose from the fact that the unmodified formalism of quantum theory is in general incapable of distinguishing an *actual* result of a measurement from all the possible results. In our model, for example, all the possible values $A_2(a_1)$, representing all the possible results $\phi_{A_2(a_1)}$ of the measurement, appear in the representation (27) of a generic state. Everett's solution begins with the simple but startling inference from (27) and Axiom 7 that the quantum formalism is in consistent with there being in general an "actual" result of a measurement in the usual sense (i.e., a single value $\phi_{A_2(\bar{a}_1)}$). Instead, he proposes that the system and apparatus observables are in general multivalued, possessing all the eigenvalues whose eigenstates appear in the representation (27). Furthermore, although nothing in the quantum formalism gives preference to any of these values over the others, the structure of (27) does pair off values of $\hat{\phi}_1(t'')$ and $\hat{\phi}_2(t'')$ with each other. Thus the value ϕ_{a_1} appears with $\phi_{A_2(a_1)}$. Each such pair of single values (or in the more general case of N subsystems, each N-tuplet of single values) would be as much as could be simultaneously specified about the whole universe, if the subsystems were kinematically independent. Everett's interpretation of (27) is "There are n_1 universes. In the a_1th universe $\hat{\phi}_1(t'')$ possesses the value ϕ_{a_1} and $\hat{\phi}_2(t'')$ possesses the value $\phi_{A_2(a_1)}$." Thus the observables are multivalued. But in a given universe, not only does $\hat{\phi}_1(t'')$ possess a single definite value, but $\hat{\phi}_2(t'')$ has correctly measured this value ("correctly," only because our model measurement was perfect). The number of universes is not reduced by further measurements, nor can an observer detect *via* repeated measurements of $\hat{\phi}_1(t'')$ and $\hat{\phi}_2(t'')$, the presence of other universes [see equation (31)]. The qualification "of $\hat{\phi}(t'')$ and $\hat{\phi}_2(t'')$" is significant and necessary, as we shall see in the following sections.

The rudimentary interpretational axiom (6) can now be omitted. Then neither the term *measurement*, nor of course the distinction between "elementary" and "complete" measurements appears in either the formalism or the interpretation of quantum theory. Unlike the C.I., the Everett interpretation can therefore be applied at all instants, not just after measurements. It gives a picture of a world (i.e., everything that exists) consisting of many coexisting universes (i.e., maximal sets observables with values) evolving approximately independently on large scales, but in intimate interaction, through interference effects, on small scales.

After a general measurement, a copy of the apparatus is present in each universe. Each copy has recorded a different result, and further measurements invariably agree with this result. Thus, since each copy began with the same (receptive) value, the deterministic evolution of the world gives rise to an irreducible unpredictability in measurements.

Yet, as I have discussed above, the outcome of measurements is not utterly unpredictable, but is *random* with a predictable probability distribution $P(a)$ [see (28)]. Everett, and later DeWitt and Graham (1973) claimed to have shown that (28) is a consequence of the Everett interpretation as outlined above, and does not have to be postulated separately. This claim is based on their discussion of sequences of measurements on identical systems, in particular on the following argument, adapted here to our perfect measurement model: Let N uncorrelated, identical system-apparatus pairs undergo the perfect measurement interaction (7). When the state $|\psi\rangle$ is decomposed in the appropriate postmeasurement basis [analogously to (27)], it will represent n_1^N universes (i.e., n_1 outcomes for each of N independent measurements). In each universe a particular distribution of outcomes happens, and this distribution can be subjected to a statistical test (e.g., the "χ^2" test) to determine whether it deviates significantly from the distribution $P(\bar{a}_1)$. In some ("maverick") universes it will, and in some it will not. Such a statistical test is also a measurement. The relevant observable is a projection operator $\hat{\mathscr{P}}$ whose eigenvalue is zero for eigenstates failing the test and one for eigenstates passing the test. Thus the state

$\hat{\mathscr{P}}|\psi\rangle$ represents some number (less than n_1^N) of universes in all of which the distribution of measurements does not differ significantly from $\mathscr{P}(\bar{a}_1)$. DeWitt and Graham then show that for any fixed statistical test $\hat{\mathscr{P}}$

$$\lim_{N\to\infty} [\hat{\mathscr{P}}|\psi\rangle - |\psi\rangle] = 0 \tag{38}$$

Thus, in the vector representing the result of an infinity of measurements, the length of the component representing "maverick" worlds is zero. Therefore the maverick universes are of measure zero (in the Hilbert space norm sense) in the set of all universes (where an infinity of measurements has been performed). Unfortunately this argument does *not* establish DeWitt and Graham's claim, for there is nothing in the formalism telling us that a set of worlds of measure zero must "occur with zero probability." Indeed, if we had been willing to identify the Hilbert space norm measure with a physical probability, then DeWitt and Graham's elaborate argument about sequences of measurements would be redundant, since the *norm* of the component $C^{a_1}|a_1, t''; A_2(a_1), t''\rangle$ in (27) is in any case $|C_{a_1}|^2$.

In order to solve this problem, I propose a slight change in the Everett interpretation:

Axiom 8. The world consists of a continuously infinite-measured set of universes.

By a "measured set" I mean a set together with a measure on that set. The interpretation of the state (27) will be that the set of universes consists of n_1 disjoint subsets, where the a_1th subset is of measure $|C_{a_1}|^2$. Each of these subsets, which I shall call a *branch*, consists of a continuous infinity of identical universes. During the model measurement, the world has initially only one branch, and is partitioned into n_1 branches. The branches play the same role as individual universes do in Everett's original version, but the probabilistic interpretation is now truly "built in." It is also sometimes claimed by proponents of the Everett interpretation that the quantum formalism (Axioms 1-7) *admits* only one interpretation [cf. DeWitt's "metatheorem," (DeWitt, 1970, 1968, 1973)]. The modification I have just introduced, small though it is, is a counterexample to that claim.

Just as in the C.I., the system-apparatus product structure and the bases $|a_1, t\rangle$, $|a_2, t\rangle$ of eigenstates of $\hat{\phi}_1(t'')$ and $\hat{\phi}_2(t'')$ are not specified, yet play a central role in the Everett interpretation of the quantum theory of our model system.

Both interpretations may conveniently be formulated so as to refer to these structures only via the preferred ordered orthonormal basis $\{|\alpha, t\rangle\}$ which they generate in the overall state space $\mathscr{H}$. I shall call $\{|\alpha, t\rangle\}$ the *interpretation basis* at time *t*. The actual ordering of the interpretation basis states is a matter of convention, which could for example be chosen so that $\alpha = 1$ corresponds via (6) to $a_1 = a_2 = 1$, $\alpha = 2$ to $a_1 = 1$, $a_2 = 2, \ldots \alpha = n_2 + 1$ to $a_1 = 2$, $a_2 = 1$, and so on.

In terms of the interpretation basis the central assertion of the C.I. may be stated thus: "At every instant t_i at which a measurement is completed the state $|\psi\rangle$ changes to an element $|\alpha, t_i\rangle$ of the interpretation basis with probability $|\langle\psi|\alpha, t_i\rangle|^2$." And the Everett interpretation says the following:

Axiom 9. At each instant t, in a proportion $|\langle\psi|\alpha, t\rangle|^2$ of all universes the value of any observable $\hat{O}$ diagonal in the interpretation basis is $\langle\alpha, t|\hat{O}|\alpha, t\rangle$.

The state $|\alpha, t\rangle$ is known as the *relative state* of the αth branch at time t.

Axioms 8 and 9, which express the Everett interpretation may be rephrased as follows: "The set of all universes is an ensemble with density operator

$$\hat{\rho}_\varepsilon(t) = \sum_\alpha |\langle\psi|\alpha, t\rangle|^2 |\alpha, t\rangle\langle\alpha, t| \text{"} \tag{39}$$

At instants of completion of perfect measurements, this expression is identical with that of the auxiliary density operator (32) which describes the ensemble version of the C.I. Thus, at such instants, Everett's many universes provide a realization of the C.I. ensemble, which we shall see in Section 7 accounts for the close similarity between the two interpretations in practice.

In view of this similarity, the reader should be warned that the role of the density operator $\hat{\rho}_\varepsilon(t)$ is utterly different in the two "interpretations." In the C.I., $\hat{\rho}_\varepsilon(t'')$ summarizes the probabilistic predictions of quantum theory about any measurements which might be made, at times greater than t'', on the world, i.e., subsystems 1 and 2 combined, by a third, kinematically independent, subsystem. The associated ensemble is a fictitious one whose properties are used to model the probabilistic properties of such measurements. Only the ensemble averages $\langle\langle\hat{O}(t)\rangle\rangle$ of observables, calculated from

$$\langle\langle\hat{O}(t)\rangle\rangle = \text{Tr}\, \hat{\rho}_\varepsilon(t'')\hat{O}(t) \qquad (t \geq t'') \tag{40}$$

have significance in the model, being equal to the statistical expectation value of the result of a measurement of $\hat{O}(t)$. Though it is possible to regard the quantum statistical ensemble, just like the Everett ensemble, as an infinite set of quantum systems in definite states, none of these states will in general be an eigenstate of $\hat{O}(t)$. Thus the result of an individual measurement of $\hat{O}(t)$ is in general realized by no property of the ensemble.

By contrast, in the Everett interpretation, $\hat{\rho}_\varepsilon(t'')$ summarizes the results of the measurement which *ended* at time t''. This was, by hypothesis, a measurement of a subsystem 1 observable by a subsystem 2 observable. The results *are* realized as properties of the Everett ensemble: The value measured in each branch is the eigenvalue corresponding to the relative state of that branch, and the probability of that value equals the measure of the branch. Observables diagonal in the interpretation basis are assigned (multiple) values, but general observables are not. The statistical properties of further measurements by a third subsystem can*not* in general be inferred from $\hat{\rho}_\varepsilon(t'')$ in the Everett interpretation. This function is performed, as ever, by the true density operator of the world

$$\hat{\rho} = |\psi\rangle\langle\psi| \tag{41}$$

via

$$\langle\langle \hat{O}(t) \rangle\rangle = \mathrm{Tr}\, \hat{\rho}\hat{O}(t) \tag{42}$$

A many-universes realization of the results of this further measurement is described neither by $\hat{\rho}$ nor by $\hat{\rho}_\varepsilon(t'')$ but by an Everett density operator for all three subsystems, whose form will depend on the *actual* measurement mode by the third subsystem.

Notice also that the properties of a quantum statistical ensemble depend upon its members being strictly independent (dynamically and kinematically) of each other whereas those of the Everett ensemble depend on interference between its members.

(40) and (42) show explicitly that the difference between the C.I. and the Everett interpretation is empirical as well as metaphysical. They make different predictions in general. In the next section we shall see why this difference is so difficult to detect in the laboratory, but in Section 8 I shall describe an experiment which would nevertheless detect it.

In this section and the previous one we have seen that there is a *lacuna* in existing formulations of both the C.I. and the Everett interpretation: In neither is the interpretation basis specified. If quantum theory is to be a universal physical theory then not only must a general specification be found for the interpretation basis for a general quantum system, but this specification must be expressible wholly within the quantum formalism. In the next section I give such a specification.

7. THE INTERPRETATION BASIS

In this section I address the problem of finding the interpretation basis for a general quantum system. My approach is heuristic, using the theory of measurement developed in the preceding sections as a tool. Later I shall indicate how other approaches would give the same answer. We shall find that at the instant of completion of a measurement, the interpretation basis is determined by the requirement that a measurement has indeed taken place. For the C.I., this restricted result solves the interpretation basis problem, though of course the other problems of the C.I. outlined in Section 5 remain. I do not attempt to solve these, and I doubt that a solution is possible. The Everett interpretation, on the other hand, suffers from no deficiency other than the lack of an interpretation basis. Since in the Everett interpretation "measurements" are not fundamentally distinguished from any other interactions, we can avoid having to append such a distinction to it if the interpretation basis construction we arrive at for measurements is applicable generally. When we have constructed it, we shall see that it is.

Consider the instant t'' of completion of the model measurement of Section 4 [equations (15), (17) or (25), (27)]. Our problem is to determine, from the kinematics and dynamics of the world at time t'', together with the fact that the measurement is completed at that instant, the interpretation basis

$$\{|\alpha, t''\rangle\} = \{|a_1, t''; a_3, t''\rangle\} \tag{43}$$

As the notation of (43) suggests, it is convenient to separate the problem into two parts: (1) The determination of the product structure delineating the subsystems 1 and 2: As discussed in Section 3, a product structure $\mathbf{L}$ is represented by an equivalence class of objects $L^{\alpha}_{s_1 s_2}$ under unitary transformations on the indices s_1 and s_2 [see (9)] satisfying the unitarity conditions (7). (2) The determination within $\mathscr{H}_1$ and $\mathscr{H}_2$ of the observables $\hat{\phi}_1(t') = \hat{\phi}_1(t'')$ and $\hat{\phi}_2(t'')$ involved in the measurement. More specifically we are really only interested in their eigenstates $\{|s_1, t''\rangle\}$ and $\{|s_2, t''\rangle\}$. From these structures, the interpretation basis can be constructed:

$$\mathscr{H} \ni |s_1, t''; s_2, t''\rangle \leftrightarrow |s_1, t''\rangle |s_2, t''\rangle \in \mathscr{H}_1 \times \mathscr{H}_2 \tag{44}$$

or in another notation

$$\mathscr{H} \ni |\alpha, t''\rangle \leftrightarrow \sum_{s_1 s_2} \mathscr{L}^{s_1 s_2}_{\alpha} |s_1, t''\rangle |s_2, t''\rangle \in \mathscr{H}_1 \times \mathscr{H}_2 \tag{45}$$

In setting up the model measurement, our first guiding principle was that if the measured observable $\hat{\phi}_1(t')$ possessed a definite value [equation (15)] then so should the "apparatus" observable $\hat{\phi}_2(t'')$. Since the measurement is nonperturbing this implies that the subsystems 1 and 2 would in that case be kinematically independent at time t''. Moreover since, by hypothesis, the dependence of $\hat{\phi}_2$ on $\hat{\phi}_1$ ceases after t'', we may infer that they would remain kinematically independent thenceforward. Physically this means that the moment of completion of the model measurement is determined by the condition that the evolution of $\hat{\phi}_2$ no longer depend on $\hat{\phi}_1$, though they may still be evolving under the trivial action of their "self-interaction" Hamiltonians $\hat{H}_1$ and $\hat{H}_2$. Thus, so far we have gained the information that in each of the states $|a_1, t''; A_2(a_1), t''\rangle$, the subsystems are dynamically independent. That is, (see Section 3), for some $\hat{H}_1$ and $\hat{H}_2$, the states $|a_1, t''; A_2(a_1), t''\rangle$ are all eigenstates of $\hat{H}-\hat{H}_1-\hat{H}_2$. In a more realistic model, where, say, imperfections allowed all the interpretation basis states to appear as final states, instead of just the n_1 accurate ones, we could repeat the above argument for every case where $|\psi\rangle$ was an interpretation basis state $|a_1, t''; a_2, t''\rangle$. Thus *there must exist a choice of* $\hat{H}_1$ *and* $\hat{H}_2$ *such that* $\hat{H}-\hat{H}_1-\hat{H}_2$ *is diagonal in the interpretation basis.* We shall see in a moment that this requirement essentially fixes the product structure **L**, but before I make this explicit, let us suppose **L** given, and use an argument based on *external* measurements to determine the preferred bases $\{|s_1, t''\rangle\}$ and $\{|s_2, t''\rangle\}$. This time we consider two subsystems in a general state

$$|\psi\rangle = \sum_{a_1 a_2} C^{a_1 a_2}|a_1, t; a_2 t\rangle \tag{46}$$

not necessarily the result of a measurement. Suppose that two more apparatuses (for which we shall fortunately not require models) measure observables $\hat{\xi}_1(t'')$ and $\hat{\xi}_2(t'')$, which are confined to subsystems 1 and 2. Now $\hat{\xi}_1(t'')$ and $\hat{\xi}_2(t'')$ are simultaneous (in a relativistic theory they need only have spacelike separation). We therefore require that the probability distribution function for the result of measuring $\hat{\xi}_1(t'')$ not depend on which observable $\hat{\xi}_2(t'')$ is chosen for the second measurement. If it did, this effect could be used to send superluminal signals, or in general, signals not carried by dynamical evolution. In terms of the Copenhagen interpretation we are requiring that it not be possible to signal by completing a measurement and triggering the discontinuous change in the global state vector. If subsystem 1 is measured just before the instant t when a measurement of subsystem 2 is completed, the probabilities of the several results are described by its density operator

$$\hat{\rho}_1(t) = \operatorname*{Tr}_{2,t} \hat{\rho} \tag{47}$$

where

$$\langle a_1 t; a_2, t|\operatorname*{Tr}_{2,t} \hat{\rho}|b_1, t; b_2 t\rangle = \sum_{\substack{\alpha,\beta \\ c_2}} \mathscr{L}^{a_1 c_2}_{\alpha}(t)\langle\alpha|\hat{\rho}|\beta\rangle\mathscr{L}^{\beta}_{b_1 c_2}(t)\delta^{a_2}_{b_2} \tag{48}$$

But if subsystem 1 were measured just after the instant t, then the state would already have undergone its random change with a statistical distribution given by $\hat{\rho}_\varepsilon(t)$, and therefore the results of the measurement on subsystem 1 will be described by the density operator

$$\operatorname*{Tr}_{2,k} \hat{\rho}_\varepsilon(t) \tag{49}$$

instead of (47). To exclude superluminal signaling it is necessary and sufficient that (49) and (47) be equal. That is,

$$\hat{\rho}_1(t) = \sum_{a_1 a_2 b_2} |c_{a_1 b_2}|^2 |a_1, t; a_2, t\rangle\langle a_1, t; a_2 t| \tag{50}$$

Hence $\hat{\rho}_1(t)$ is diagonal in the interpretation basis. And thus *the interpretation basis at time t is the basis of eigenstates of* $\hat{\rho}_1(t)$ and (by the same argument) of $\hat{\rho}_2(t) = \mathrm{Tr}_{1,t}\,\hat{\rho}$.

Our earlier requirement that $\hat{H} - \hat{H}_1 - \hat{H}_2$ be diagonal in the interpretation basis may therefore be written

$$[\hat{H} - \hat{H}_1 - \hat{H}_2, \hat{\rho}_1(t) \times \hat{\rho}_2(t)] = 0 \qquad \text{for some } \hat{H}_1, \hat{H}_2 \tag{51}$$

The "self-interaction" Hamiltonians $\hat{H}_1$ and $\hat{H}_2$ may be eliminated from this equation by substituting for the partial traces

$$\begin{aligned} \operatorname*{Tr}_{1,t}[\hat{H}, \hat{\rho}_1 \times \hat{\rho}_2] &= [\hat{H}_2, \hat{\rho}_2] \\ \operatorname*{Tr}_{2,t}[\hat{H}, \hat{\rho}_1 \times \hat{\rho}_2] &= [\hat{H}_1, \hat{\rho}_1] \end{aligned} \tag{52}$$

Hence

$$[\hat{H}, \hat{\rho}_1 \times \hat{\rho}_2] - \hat{\rho}_1 \times \operatorname*{Tr}_{1,t}[\hat{H}, \hat{\rho}_1 \times \hat{\rho}_2] - \operatorname*{Tr}_{2,t}[\hat{H}, \hat{\rho}_1 \times \hat{\rho}_2] \times \hat{\rho}_2 = 0 \tag{53}$$

(53) must of course be regarded as an equation for the product structure **L** implicit in $\hat{\rho}_1$, $\hat{\rho}_2$, $\mathrm{Tr}_{1,t}$ and $\mathrm{Tr}_{2,t}$. In this paper I shall not pursue the problem of the existence or uniqueness of solutions of (53) beyond the following crude but encouraging argument: First let us calculate the number of independent components of **L**. Since $\mathscr{L}^{\alpha}_{a_1 a_2}$ is unitary it has $n_1^2 n_2^2$ independent (real) components, where n_1 and n_2 are the dimensions of $\mathscr{H}_1$ and $\mathscr{H}_2$. But **L** has fewer components because unitary transformations within $\mathscr{H}_1$ and $\mathscr{H}_2$ do not affect it. The set of such transformations is parametrized by $n_1^2 + n_2^2 - 1$ components. The "-1" is there because the trivial unitary transformation $\mathscr{L}^{\alpha}_{a_1 a_2} \to e^{ip}\mathscr{L}^{\alpha}_{a_1 a_2}$ with p real has representations both in $\mathscr{H}_1$ and $\mathscr{H}_2$. Hence **L** has $n_1^2 n_2^2 - n_1^2 - n_2^2 + 1$ independent components. Now (53) has, *prima facie*, $n_1^2 n_2^2$ complex components, but it is anti-Hermitian and traceless under both $\mathrm{Tr}_{1,t}$ and $\mathrm{Tr}_{2,t}$. This leaves $n_1^2 n_2^2 - n_1^2 - n_2^2 + 1$ independent real equations in (53), the same as the number of independent components in the product structure.

Equations (50) and (53) formally define the interpretation basis, given $|\psi\rangle$ and $\hat{H}$, at any instant. Thus they manifestly provide the structure missing from existing interpretations of quantum theory, and they construct it *almost* entirely within the quantum formalism. But we have not quite finished yet.

There were still two pieces of *a priori* metaphysics (i.e., information not contained in $|\psi\rangle$ and $\hat{H}$) which we supplied while constructing (50) and (53) for our model system. One was that the world was divided into precisely two subsystems, and the other was the dimensionalities, n_1 and n_2, of the subsystems' state spaces (though the value of their product, $n_1 n_2 = n$ *is* an objective property of the world). Amusingly, this ambiguity depends, in the case where n is finite, on the prime factorization of n. If n is prime there can be no subsystem of the world and the above construction of the interpretation basis will not work. It would do no harm to prohibit such quantum theories since measurement, and hence experimental testing of the theory, cannot take place in the worlds they describe. If n is a product of two primes, then (50) and (53) give a unique interpretation basis without any need for the *a priori* metaphysics. However, the observed richness of the world presumably indicates that if n is finite, then it is highly composite.

When the number of subsystems is greater than 2, the interpretation basis is no longer the direct product of bases confined to the subsystems' state spaces. This is illustrated by the following Stern-Gerlach example: A spin-$\frac{1}{2}$ silver atom in a generic state passes through an ideal Stern-Gerlach apparatus where its spin component $\mathbf{n} \cdot \hat{s}$ is perfectly measured by its emergence angle $\hat{\theta}_2$ (eigenvalues $\pm\theta_2$, say). A second apparatus, placed so that both $(\pm\theta_2)$ trajectories enter it, is so constructed that it has different effects on atoms on the two trajectories, but so that both effects are perfect measurements. For atoms on the $+\theta_2$ trajectory, the emergence angle $\hat{\theta}_3$ (eigenvalues $\pm\theta_3$) measures $\mathbf{l} \cdot \hat{s}$, and for atoms on the $-\theta_2$ trajectory $\hat{\theta}_3$ measures $\mathbf{m} \cdot \hat{s}$, where $\mathbf{l}$, $\mathbf{m}$ and $\mathbf{n}$ are not coplanar. Then the interpretation basis after the measurements must contain the states

$$\begin{aligned} &|\mathbf{l} \cdot \mathbf{s} = +\tfrac{1}{2}\hbar\rangle|+\theta_2\rangle|+\theta_3\rangle \\ &|\mathbf{l} \cdot \mathbf{s} = -\tfrac{1}{2}\hbar\rangle|+\theta_2\rangle|-\theta_3\rangle \\ &|\mathbf{m} \cdot \mathbf{s} = +\tfrac{1}{2}\hbar\rangle|-\theta_2\rangle|+\theta_3\rangle \\ &|\mathbf{m} \cdot \mathbf{s}| = -\tfrac{1}{2}\hbar\rangle|-\theta_2\rangle|-\theta_3\rangle \end{aligned} \tag{54}$$

which, since the eigenstates of $\mathbf{l} \cdot \hat{s}$ do not coincide with those of $\mathbf{m} \cdot \hat{s}$, cannot be part of a product basis.

However, the interpretation basis can be constructed, when the state is a linear superposition of (54), as follows. Consider the world as being split into *two* subsystems first, one with a four-dimensional state space, which will turn out to be $\mathcal{H}_1 \times \mathcal{H}_2$, and one with a two-dimensional state space $\mathcal{H}_3$. For any given such product structure (to be determined later), we can calculate $\hat{\rho}_3$, and precisely analogous requirements to those for the 2-subsystem case show that the interpretation basis states be eigenstates of

$\hat{\rho}_3$: This gives the $|\pm\theta_3\rangle$ part of (54). The reason why we start by splitting off subsystem 3 rather than 2 or 1 is not yet evident. Notice that it makes a considerable difference: Starting with subsystem 1 would give the "wrong" answer [i.e., not (54)]. I shall return to this problem below.

More complicated branching schemes are possible for three than for two subsystems. In particular there is now more than one branch present for each eigenvalue θ_3. The structure of these branches is again determined by the 2-subsystem arguments, now required to hold for each relative state $|\pm\theta_3\rangle$ of subsystem 3. The interpretation bases for the two "auxiliary worlds" with states

$$|\psi, \pm\theta_3\rangle = \langle\pm\theta_3|\psi\rangle \tag{55}$$

and Hamiltonians

$$\hat{H}_{\pm\theta_3} = \langle\pm\theta_3|\hat{H}|\pm\theta_3\rangle \tag{56}$$

are determined from (50) and (53). The interpretation basis is now determined, except for the product structure which we left undetermined at the beginning. That, as always, is fixed by the requirement that $\hat{H} - \hat{H}_3 - \hat{H}_{1,2}$ be diagonal in the interpretation basis for some $\hat{H}_3$ and $\hat{H}_{1,2}$, when $\hat{H}_{1,2}$ is confined to the combined 1, 2 subsystem, though this requirement is no longer expressible in the simple form (53).

The construction of the interpretation basis in this 3-subsystems case admits, I suggest, only one generalization to the case where the dimensionality of $\mathcal{H}$ has arbitrarily (but finitely) many factors. This is summed up by the following recursive algorithm:

7.1. Construction of the Interpretation Basis for a General Quantum System with a Finite-Dimensional State Space

Given $\mathcal{H}$, $|\psi\rangle$ and $\hat{H}(t)$:

Step 1. Let n be the dimensionality of $\mathcal{H}$. The case where n is prime is not permitted. If n is a product of two primes then equations (50) and (53) determine the interpretation basis. Otherwise,

Step 2. Let $n = n_1 n_2$ where n_1 is prime and n_2 is composite. (*Which* prime factor is taken as n_1 is discussed below.) Take an arbitrary product structure $\mathsf{L}(t)$, dividing the world at each instant t into an n_1- and an n_2-dimensional subsystem. For each eigenstate $|a, t; \mathsf{L}\rangle$ of $\mathrm{Tr}_{2,t}|\psi\rangle\langle\psi|$, apply Step 1 to subsystem 2, considered as an "auxiliary world" with state $\langle a_1, t; \mathsf{L}|\psi\rangle$ and Hamiltonian $\langle a_1, t; \mathsf{L}|\hat{H}(t)|a_1, t; \mathsf{L}\rangle$. Let the interpretation basis thus obtained for subsystem 2 be $\{|\alpha, t, a_1, \mathsf{L}\rangle\}$ $(1 \leq \alpha \leq n_2)$. The full interpretation basis is $\{|a_1, t; \mathsf{L}\rangle|\alpha; t, a_1, \mathsf{L}\rangle\}$.

Step 3. $\mathsf{L}(t)$ itself is now determined by requiring that $\hat{H}(t)-\hat{H}_1(t)-\hat{H}_2(t)$ be diagonal in the interpretation basis for some choice of operators $\hat{H}_1(t)$ and $\hat{H}_2(t)$ confined to subsystems 1 and 2, respectively. □

In this general case, as in the case of two subsystems, the term "interpretation basis" is used to denote an *ordered* basis. The ordering is chosen so as to represent, according to some convention which we need not specify, the preferred product structures at each level. The information thus represented I shall call the "branching structure."

At two points in the above algorithm degeneracies (in $\mathrm{Tr}_{2,t}|\psi\rangle\langle\psi|$ and $\hat{H}(t)-\hat{H}_1(t)-\hat{H}_2(t)$) seem to spoil the uniqueness of the specification of the interpretation basis. Since there is strong reason to suppose (see Section 9) that the real world is strongly degenerate in these quantities, this non-uniqueness cannot be argued away as a problem of "measure zero." It is nevertheless trivial for another reason: Such degeneracies occur precisely when there is nothing in the world to distinguish between the different interpretations arising from the equivalence class of interpretation bases generated by the algorithm. The interpretations differ in form only. For example, consider two spin-$\frac{1}{2}$ subsystems in the state

$$|\psi\rangle=\frac{1}{\sqrt{2}}(|\uparrow\rangle|\downarrow\rangle+|\downarrow\rangle|\uparrow\rangle) \tag{57}$$

This is equal to

$$|\psi\rangle=\frac{1}{\sqrt{2}}(|\rightarrow\rangle|\leftarrow\rangle+|\leftarrow\rangle|\rightarrow\rangle) \tag{58}$$

where with the obvious notation $|\uparrow\rangle$ denotes the spin "up" eigenstate of the "vertical" spin operator and so forth. The identity in form between (57) and (58) is reflected in a degeneracy in the density operators $\mathrm{Tr}_1|\psi\rangle\langle\psi|$ and $\mathrm{Tr}_2|\psi\rangle\langle\psi|$ and hence in an ambiguity in the interpretation basis. Alternative interpretations of a world in the state $|\psi\rangle$ are apparently possible: [from (57)] "In the first branch, system 1 has spin up and system 2 has spin down ... etc." or [from (58)] "In the first branch system 1 has spin North and system 2 has spin South ... etc." The possibility of more than one interpretation is not consistent with the principle of realism which we have been implementing everywhere. But of course the terms "up," "down," "North," and "South" have no invariant physical meaning in the simple world of (57). They would acquire meaning only if the degeneracy were broken by the dynamical evolution.

The other ambiguity in the algorithm is more real. Locally, the order in which the subsystems are split off during the construction of the interpretation basis is determined by continuity: The interpretation basis and the associated branching structure must be continuous in time. But for a general quantum system this still allows many distinct choices, one for each choice of the branching structure in the remote past. It is always assumed in physics that sufficiently far in the past, the preferred subsystems of the world were kinematically independent. I have assumed it, for example, in all the model measurements described in this paper. The justification of this assumption comes from considerations, such as causality and the second law of thermodynamics, which have not yet been incorporated into quantum theory at a fundamental level (but see Section 9). But if the assumption is true, then there is only one possible branching structure in the remote past, and the interpretation basis is thereby uniquely determined at all times.

7.2. Other Properties of the Interpretation Basis

The physical considerations from which we have just constructed the interpretation basis are not the only ones from which it may be determined.

One illuminating approach is to regard the C.I. as a "classical limit" of the Everett interpretation. This limit is not that of the usual correspondence principle (though there may be some connection: see Section 9), but is the limit where one is unable to measure the kinematical dependence between subsystems—specifically between a system and its observer: After a (complete) measurement, the probability distribution function for further measurements on the same subsystem (subsystem 1, in the model), calculated from $\mathrm{Tr}_2\, \hat{\rho}$ in the Everett interpretation, must be identical with that calculated from the auxiliary density operator $\mathrm{Tr}_2\, \hat{\rho}_\varepsilon$ which represents the mixed case predicted by the C.I. Equation (50) follows.

Unfortunately, I have not been able to extend this "correspondence principle" to derive equation (53) for the product structure.

Another way of looking at the conditions (50) and (53) is that they are the requirements that the Everett branches cannot communicate with each other. Equation (50) says that after a measurement, further measurements on the same subsystem cannot reveal information about all the branches, as contained in the full density operator $\hat{\rho}$. (53) says that the time evolution of each branch is instantaneously independent of every other branch.

From equations (47) and (50) we may deduce an algebraic property of the interpretation basis (d'Espagnat, private communication). Take $n_1 = n_2$ for simplicity. We have

$$\sum_{c_2} c_{a_1 c_2} c^{b_1 c_2} = \sum_{c_2} |c_{a_1 c_2}|^2 \delta^{b_1}_{a_1} \tag{59}$$

and from the analogous expression to (50) for $\mathrm{Tr}_1 \hat{\rho}$,

$$\sum_{c_1} c_{c_1 a_2} a^{c_1 b_2} = \sum_{c_1} |c_{c_1 a_2}|^2 \delta^{b_2}_{a_2} \tag{60}$$

Multiplying (60) by $c_{a_1b_2}$, summing over b_2, and using (59) we obtain

$$c_{a_1a_2}[\sum_{b_2}|c_{a_1b_2}|^2-\sum_{b_1}|c_{b_1a_2}|^2]=0 \tag{61}$$

whence we deduce that in the interpretation basis, $c_{a_1a_2}$ has *qua* matrix no more than one nonzero element per row or column. Therefore it is possible to relabel the interpretation basis states $|a_1, t\rangle$ and $|a_2, t\rangle$ so that $c_{a_1a_2}$ takes precisely the form of the state (27) following a perfect measurement. This is known as the Schmidt normal form (Schmidt, 1907) of the vector $|\psi\rangle$ with respect to the given product structure.

The condition (50) on the interpretation basis was first obtained by Zeh (1973; see also Schrödinger, 1935) using physical arguments different from those of this paper.

There is therefore a sense in which every interaction is a perfect measurement of something, though this fact must be considered with caution for several reasons. Firstly, the measured observable is the perturbed one $\hat{\phi}_1(t'')$ (in the notation of Section 3), from whose measurement in general nothing can be inferred about observables $\hat{\phi}_1(t')$ at an earlier time. Secondly, if the "system" and "apparatus" were not kinematically independent at an earlier time t', before the interaction, then the interaction cannot be interpreted as having effected a measurement in the usual sense: The "measurement arrow of time" is missing (see Section 9). Thirdly, the accuracy or otherwise of a measurement depends, as I have mentioned before, not only upon the constitution of the apparatus, but also upon the intention of the observer: An apparatus which measures $\hat{\phi}_1$ perfectly may be considered inaccurate by an observer who built it as a $\hat{\xi}_1$ meter.

The subjective overtones of the term "intention," and the result (61) make a short discussion of *imperfect* measurements advisable here. If imperfection is in the eye of the observer, it follows that a model of an imperfect measurement must include a model of this "eye," incorporating the observer's intention in an objective way (cf. DeWitt, 1968, 1973). The simplest model has three subsystems. Subsystems 1 and 2 are the same as in the model of Sections 3 and 4, though we need not now specify what interactions they have undergone. We now know that in any case $|\psi\rangle$ will take the perfect measurement form (27) in the interpretation basis. This time, however, subsystem 3 will act as a second apparatus, measuring the accuracy with which a given subsystem-1 observable, say $\hat{\xi}_1(t')$, was measured by a given subsystem 2 observable $\hat{\xi}_2(t'')$. To avoid complicating matters let us assume that this second measurement is perfect. There are many possible measures of accuracy. In a simple case, subsystem-3 would have an n_1n_2-dimensional state space and might measure the 1, 2 observable

$$\hat{A} = \sum_{a_2 = A_2(a_1)} |a_1, t'; a_2, t''\rangle\langle a_1, t'; a_2, t''| - \sum_{a_2 \neq A_2(a_1)} |a_1, t'; a_2, t''\rangle\langle a_1, t'; a_2, t''| \tag{62}$$

The value $\langle \hat{A} \rangle = +1$ indicates "accurate," anything less than 1, "inaccurate," and -1, "perfectly inaccurate," where $|a_1, t'\rangle$ and $|a_2, t''\rangle$ are the eigenstates of $\hat{\xi}_1(t')$ and $\hat{\xi}_2(t'')$, respectively. During the course of the second measurement, the interpretation basis would be forced to change continuously from eigenstates of $\hat{\phi}_1(t'')$ and $\hat{\phi}_2(t'')$ (say) to eigenstates of $\hat{\xi}_1(t')$ and $\hat{\xi}_2(t'')$. Thus the intention of the observer is reflected in objective properties of subsystem 3 and its interactions with subsystems 1 and 2, and can be interpreted *via* the interpretation basis in the usual way.

7.3. Deficiencies in Above Approach

The methods of this paper, leading up to the algorithm presented earlier in this section, solve the problem of the interpretation basis for quantum theories with finite-dimensional state spaces. If the same thing could be done for field theories then, at least for those who find Everett's interpretation acceptable, the "problem of measurement" and the problem of the interpretation of quantum theory in general, would be solved. Quantum theory could be regarded without reservation as a universal physical theory.

However, it seems difficult to extend the methods to field theories, since they appear to depend heavily on properties, such as unique prime factorizability, which are characteristic of the integers.

In order to obtain a unique interpretation, I was forced to assume something about the state $|\psi\rangle$ of the world: That in the remote past, the interpretation subsystems were kinematically independent. This assumption appeals to thermodynamical principles outside quantum theory. Optimistically, (see Section 9), this might suggest that the principles of thermodynamics are derivable from a fuller quantum theory—but by the same token one would infer that existing quantum theory is not a complete theory of nature.

Finally, the quantum theory presented in this paper is necessarily nonrelativistic. DeWitt's change from the Schrödinger to the Heisenberg picture, followed here, eliminated from the Everett interpretation some of its dependence on the absolute time t and is a necessary first step toward the interpretation of relativistic theories. But until the remaining steps are taken we shall not know whether the machinery set up here is sufficient.

8. A THOUGHT EXPERIMENT

I have pointed out several times in the preceding sections that the C.I. and the Everett "interpretations" are really different formalisms for quantum theory—in effect, different physical theories. It is usually claimed that although their assertions as to the nature of objective facts are radically different, the two "interpretations" agree about all subjective experiences of observers, and cannot therefore be distinguished experimentally. That this claim is false is shown by equations (40) and (42) which summarize the different predictions of the two "interpretations" concerning further measurements on subsystems one of which has completed a measurement on the other.

Now the origin of the difference between the two predictions lies in the claim of the C.I. that quantum systems undergo an anomalous time evolution, namely, a change in their Heisenberg states, at some late stage in every complete chain of measurements. Determining experimentally, by means of interference experiments, that such a change has not yet occurred at an intermediate point in measurement chain does not, as we have seen, refute the C.I. It merely demonstrates that any change must occur at a later time. For this reason, any experiment to distinguish between the "interpretations" necessarily involves measurements by the observer on himself (or equivalently, by one observer on another). And therefore the construction of a thought experiment (i.e., an experiment whose execution is at present impractical, but in principle permitted by physical laws) for this purpose necessarily involves the description of a model observer. Fortunately, we shall find that very few details of the internal constitution of the observer need to be specified in the model, the main requirement on it being that it be a subsystem of the world and obey quantum theory (in one version or the other, to be determined). This requirement is certainly satisfied if quantum theory is a universal theory. Thus, the idealizations I shall need to make in the following description of the thought experiment come mainly at the level of the *elementary* measurements at the beginning of the measurement chains; I shall assume that they are perfect measurements (perfectly accurate, nonperturbing, and dynamically and kinematically independent of the outside world, as in Section 3). The discussion would be greatly complicated, but its conclusion unaltered, by the relaxation of these idealizations. The only other important assumption is the usual one in measurement theory, that the dynamics of the apparatuses, including in this case those which measure the observer, are freely specifiable. This assumption is not very realistic, which is why this is a thought experiment. But it is a legitimate thought experiment, since there is no known or conjectured physical principle, within quantum theory or outside it, which would prohibit the couplings it calls for.

Since it will be testing statistical predictions, as usual in quantum theory, we must think of the thought experiment as being performed sufficiently many times for the results to be statistically significant.

Four subsystems are involved in this experiment. The first one is a spin-$\frac{1}{2}$ atom (so $n_1 = 2$). This passes through a Stern–Gerlach apparatus in such a way that the two exit trajectories, corresponding to spins North and South, pass over subsystems 2 and 3. These are also spin-$\frac{1}{2}$ atoms which represent part of the sensitive "sense organ" of the observer (so $n_2 = n_3 = 2$). Their receptive states are "spin down" ($|\downarrow_2\rangle$ and $|\downarrow_3\rangle$), and the coupling is such that if the subsystem 1 atom describes the North trajectory, passing over atom-2, then atom-2's spin flips with certainty to "spin up," while atom-3's spin remains unchanged. Similarly, atom-3 flips if atom-1 describes the South trajectory. After passing through the "sense organ," each trajectory enters a storage ring designed to keep any atom entering it in orbit for ever. Subsystem 4 is the observer, or the observer's brain, so perhaps n_4 is $(10^{10})^{10}$ or more—it will not matter.

In this description I am ignoring three classes of subsystems that would be present in reality: (1) The "rest of the world," which I am assuming remains kinematically and dynamically independent of the subsystems of interest, but which would in reality introduce perturbations which would have to be taken into account. (2) Items of apparatus such as the Stern-Gerlach magnets and the storage rings, and whatever agency causes the Hamiltonian of the system to change in the complicated (but fixed) way to be described below. These are of course quantum systems but I am assuming that they are set up so that their effect on subsystems 1 to 4 is describable by an effective Hamiltonian $\hat{H}(t)$ restricted to subsystems 1 to 4. (3) Subsystems, such as the position of atom 1, which have been assumed to be in such perfect correlation with one of the subsystems 1 to 4 (in this case, with the spin of atom 1) that their separate description would be superfluous.

Initially, at time t', all four subsystems are kinematically independent. The sense organ is receptive. Atom-1 is in a state

$$|\uparrow_1, t'\rangle = \frac{1}{\sqrt{2}}(|N_1, t'\rangle + |S_1, t'\rangle) \qquad (63)$$

which is not an eigenstate of the observable which is about to be measured. The observer is also in some definite state, say, $|O_4, t'\rangle$, whose exact properties we shall not require. Nor will it be necessary that this state be exactly the same on each run of the experiment, only that it have whatever property the C.I. may designate as necessary for subsystem 4 to be capable of completing a measurement. Thus

$$|\psi\rangle = |\uparrow_1, t'\rangle|\downarrow_2, t'\rangle|\downarrow_3, t'\rangle|O_4, t'\rangle \qquad (64)$$

From what we have said, this may be written in terms of interpretation basis states at a time t'', after the atom has passed through the sense organ, as

$$
\begin{aligned}
|\psi\rangle = {} & \frac{1}{\sqrt{2}}\,|N_1, t''\rangle|\uparrow_2, t''\rangle|\downarrow_3, t''\rangle|\bar{O}_4, t''\rangle \\
& + \frac{1}{\sqrt{2}}\,|S_1, t''\rangle|\downarrow_2, t''\rangle|\uparrow_3, t''\rangle\|\bar{O}_4, t''\rangle
\end{aligned} \tag{65}
$$

where

$$|\bar{O}_4, t''\rangle = |O_4, t'\rangle \tag{66}$$

Next, between times t'' and t''', the measurement is completed. Subsystem-4 has no direct access to the atom-1, which is by now in the storage rings, but the information as to whether atom-1's spin is N or S is recorded in the sense organ. For simplicity, (my conclusions do not depend on this assumption) I shall assume that the measurement chain has the property that every interaction between subsystems 2 and 3 and subsystem 4 is a nonperturbing measurement. After the completion of the measurement, the observer records (in his memory, or in his notebook if necessary)—not the value "N" or "S" of the spin, but only *whether or not he knows this value.* He may write "I, Professor X, F.R.S., hereby certify that at time t''' I have determined whether the value of the North component of the spin of atom-1 is $+\frac{1}{2}\hbar$ or $-\frac{1}{2}\hbar$. At this moment I am contemplating in my own mind one, and only one of those two values. In order to facilitate the second part of this experiment, I shall not reveal which one." This constitutes a record of the *completion* of the measurement, a record which, we shall see, need not be destroyed by a subsequent interference experiment.

Now the honest recording of such a statement involves measurements made by subsystems of subsystem 4 on each other, reducing ultimately to measurements on subsystems 2 and 3. The relevant observable is the projection operator

$$\hat{P}(t) = |\uparrow_2, t\rangle|\downarrow_3, t\rangle\langle\uparrow_2, t|\langle\downarrow_3, t| + |\downarrow_2, t\rangle|\uparrow_3, t\rangle\langle\downarrow_2, t|\langle\uparrow_3, t| \tag{67}$$

which has the eigenvalue 0 when the sense organ is still receptive (i.e., when the spin of atom-1 is not yet known), and +1 when the sense organ has recorded one of the two values (which is then known). Strictly speaking, the interpretations "spin known" and "spin unknown" do not belong to $\hat{P}(t''')$ but to the analog of $\hat{P}(t''')$ which acts on an image of subsystems 2 and 3 some way up the measurement chain. But, insofar as the observer is supposed to have an accurate knowledge of (i.e., to make accurate measurements of) observables inside himself, we introduce no error with this loose nomenclature. When measurements of $\hat{P}(t)$ are made by subsystem-4 simultaneously with measurements of the N-S spin itself [here again, we really mean an *image* of that spin (see below)], the interpretation product structure is not the obvious one for subsystems 2 and 3. Instead, it is the product structure generated by $\hat{P}(t)$ and, say

$$\hat{Q}(t)=\tfrac{1}{2}\hbar|\uparrow_2, t\rangle|\downarrow_3, t\rangle\langle\uparrow_2, t|\langle\downarrow_3, t|-\tfrac{1}{2}\hbar|\downarrow_2, t\rangle|\uparrow_3, t\rangle+\langle\downarrow_2, t|\langle\uparrow_3, t|$$
$$+\tfrac{1}{2}\hbar|\downarrow_2, t\rangle|\downarrow_3, t\rangle\langle\uparrow_2, t|-\tfrac{1}{2}\hbar|\uparrow_2, t\rangle|\uparrow_3, t\rangle\langle\uparrow_2, t|\langle\uparrow_3, t| \quad (68)$$

after time t'', $\hat{Q}(t)$ "holds an image" of the N/S spin of atom-1. We may separate the Hamiltonian $\hat{H}(t)$ which generates the dynamical evolution between times t'' and t''' into three parts: $\hat{H}_P(t)$, representing nonperturbing measurements of subsystem P by subsystem 4, $\hat{H}_Q(t)$, representing nonperturbing measurements of subsystem Q by subsystem-4, and $\hat{H}_4(t)$, representing the self-interactions of subsystem 4. (For simplicity, assume that subsystems 1, 2, and 3 ($\equiv 1$, P, and Q) have no self-interactions.) Now at time t''', after the completion of the measurement, the two interpretations differ about the state of the world. According to Everett,

$$|\psi\rangle=\frac{1}{\sqrt{2}}|N_1, t'''\rangle|+1_P, t'''\rangle|+\tfrac{1}{2}\hbar_Q, t'''\rangle|\text{“knows } N\text{”}, t'''\rangle$$
$$+\frac{1}{\sqrt{2}}|S_1, t'''\rangle|+1_P, t'''\rangle|-\tfrac{1}{2}\hbar_Q, t'''\rangle|\text{“knows } S\text{”}, t'''\rangle \quad (69)$$

whereas according to the C.I. it has changed, with equal probability, to either

$$|\psi(t''')\rangle=|N_1, t'''\rangle|+1_P, t'''\rangle|+\tfrac{1}{2}\hbar_Q, t'''\rangle|\text{“knows } N\text{”}, t'''\rangle$$

or (70)

$$|\psi(t''')\rangle=|S_1, t'''\rangle|+1_P, t'''\rangle|-\tfrac{1}{2}\hbar_Q, t'''\rangle|\text{“knows } S\text{”}, t'''\rangle$$

An interference experiment is now performed as follows. Between the times t''' and t'''', where $t''''-t'''=t'''-t''$, the Hamiltonian is set to

$$\hat{H}(t)=-\hat{H}_Q(2t_3-t)-\hat{H}_4(2t_3-t)+\hat{H}_P(t) \qquad (t'''<t<t'''') \quad (71)$$

$\hat{H}_P(t)$ may be any desired operator confined to the subsystems P and 4, but for simplicity we assume again that it does not perturb $\hat{P}$. (71) undoes *part* of the dynamical evolution that happened during the completion of the measurement: All the system-4 observables which had, by the time t''', come to depend on $\hat{Q}(t'')$ but not on $\hat{P}(t'')$, such as records of the N-S spin, are at t'''' restored to their t'' values. All other system-4 observables, in particular those recording measurements of $\hat{P}(t'')$, are not restored. But in particular, *the record that the N–S value of the spin was known to the observer at time t''' is preserved.*

At this point, t'''', according to the Everett interpretation, all copies of the observer are once again identical though they had been different in two branches at time t''' (69):

$$|\psi\rangle = \left\{ \frac{1}{\sqrt{2}} |N_1, t''''\rangle | +1_P, t''''\rangle | +\tfrac{1}{2}\hbar_Q, t''''\rangle \right.$$
$$\left. + \frac{1}{\sqrt{2}} |S_1, t''''\rangle | +1_P, t''''\rangle | -\tfrac{1}{2}\hbar_Q, t''''\rangle | \right\}$$
$$\times |\text{“knew } N \text{ or } S\text{”}, t''\rangle \tag{72}$$

According to the C.I., only one of these terms is present. Next the Hamiltonian for subsystems 1, 2, and 3 is set to the negative of what it was between times t' and t''. Subsystem 4 is isolated from the others, though its self-interaction, or indeed, interaction with the "outside world," is freely specifiable. The fields in the storage rings and in the Stern–Gerlach apparatus are also reversed, so that the atom eventually emerges from the entrance to the apparatus, having tripped the relevant "sense organ" atom back into its receptive state on the way.

According to Everett, it will now be in its original state $|\uparrow_1, t'''''\rangle = |\uparrow_1, t'\rangle$, whereas the C.I. would have it in either a N_1 or S_1 state, as measured. Now, finally, the empirical difference between the "interpretations" has become accessible. All we need is a further Stern–Gerlach apparatus to measure the "up" component of the spin of atom-1. According to the C.I., the values $\pm\frac{1}{2}\hbar$ will be observed at random with equal probability. According to Everett, the value $+\frac{1}{2}\hbar$ will be observed with certainty.

8.1. Further Comments on the Thought Experiment

The actual performance of this experiment, involving as it does fairly detailed adjustments inside an observer's brain and sense organs, is far beyond present-day technology, but perhaps not quite as far as it might seem at first sight. Sooner or later (Turing, 1950) there will be machines capable of independent thought comparable in every way to that of human beings. One of them could no doubt be persuaded to take part in this experiment. Presumably its internal workings will be electronic, rather than biological, and its "self-interaction" Hamiltonian ($\hat{H}_4$) will be known to the designers. Extra apparatuses would have to be installed to give it the requisite sense organ, and to allow the total Hamiltonian to be temporarily altered when necessary. Sufficient coherence for the interference effects to be preserved will be possible if, for example, the information in the sense organ, the memory, and all other affected parts of the observer is stored in sufficiently microscopic finite-state components, thermally isolated from the outside world. Another possibility might be to replace all the components by logically equivalent systems of currents in superconductors.

As I have already remarked, the observer need not be prepared in a predetermined state (this might be thought to prejudice his status as an autonomous thinking being, in versions of the C.I. where such considerations count), though his sense organ must be. His only preparation need be a familiarity with the purpose of the experiment. Observe also that although part of his memory, including memories of some of his own thoughts, is necessarily erased during the experiment, the important memories and thought processes (those originating with his measurement of $\hat{P}$) are, by construction, untouched.

Everett (1957) once compared critics of his interpretation who said that they "did not feel themselves split" with opponents of Copernicus who did not feel the Earth move since, he said, the laws of quantum theory *predict* that we do not "feel other branches" just as classical dynamics predicts that we do not feel the Earth move. Yet if our senses were fine enough, we could see stellar parallaxes and feel Coriolis forces and thus "feel" the motion of the Earth. Similarly, this experiment allows the observer to "feel" himself split into two branches: The interference phenomenon seen by our observer at the end of the experiment requires the presence of both spin values, though he accurately remembers having known at a previous time that only one of them was present. He must infer that there was more than one copy of himself (and the atom) in existence at that time, and that these copies merged to form his present self.

9. THERMODYNAMICS OF SINGLE SYSTEMS

Subsystems of quantum systems are described by density operators rather than states. The theory of such subsystems, ("open systems") where one strictly ignores the rest of the world, is formally identical with the theory of ensembles of closed quantum systems, i.e., quantum statistical mechanics. This formal identity is very useful to anyone studying the foundations of thermodynamics from the quantum point of view (Davies, 1976; d'Espagnat, 1976). The usual starting point is the proposition that since quantum theory makes only statistical predictions, it is in fact a statistical theory (i.e., a theory of ensembles). This *non sequitur* is unfortunate from the point of view of physics because no ensembles are known in nature. The "ordinary" quantum formalism (as in Section 2) is thus viewed as an approximation scheme, applicable only when the ensemble is a pure or nearly pure case.

In line with the general philosophy of this paper, I take the opposite point of view: Quantum theory is an objective theory of single closed systems in pure states. Some large quantum systems in some states behave approximately like ensembles; moreover, subsystems of quantum systems behave in some ways like ensembles. Thus the ensemble approximation is often a good one for the behavior of certain real physical systems. But it is never exact. Why is this approximation so often so good? What is the origin of the second law of thermodynamics? I conjecture that the answers to these questions lie entirely in the quantum theory of single closed systems. Unfortunately these answers will not be accessible until the results of Section 7 have been extended to field theories and to relativity: Only then could they be applied to a quantum cosmological model bearing the remotest resemblance to reality. However, certain features of the finite-dimensional state space theory developed in this paper already suggest how quantum theory could provide foundations for thermodynamics.

We saw in Section 6 that the "many universes" of a single quantum world may for some purposes be regarded as a quantum statistical ensemble whose density operator is not that of a pure case. With the help of the interpretation basis it is possible to define an entropy for single closed quantum systems without recourse to ensembles or coarse graining:

$$S(t) = -\mathrm{Tr}\,\hat{\rho}_\varepsilon(t) \ln \hat{\rho}_\varepsilon(t) \tag{73}$$

$$= -\sum_\alpha |\langle\psi|\alpha, t\rangle|^2 \ln |\langle\psi|\alpha, t\rangle|^2 \tag{74}$$

$S(t)$ is nonnegative definite; moreover the assumption that the interpretation subsystems were all kinematically independent in the remote past, which we needed in Section 7 to make the interpretation basis well defined, is the same as the assumption that $S(t)$ was zero then. $S(t)$ is not necessarily an increasing function of time: there can be no such function in a time reversal invariant theory such as quantum theory. It tends to increase whenever hitherto kinematically independent subsystems begin to interact. In general its motion depends on the dynamical evolution of the world and would have different properties in different cosmological models. However, in every case there will be an open time interval T immediately following the big bang during which $S(t)$ increases. I conjecture that generically in realistic closed quantum cosmological models T exceeds the expected lifetime of the universe. A demonstration of this conjecture (which, I repeat, must await a generalization of the ideas of this paper to quantum field theory) would constitute a derivation of the second law of thermodynamics as a theorem of quantum theory (including the interpretation basis). The arrow of time is built into quantum theory *via* the interpretation basis, specifically *via* the branching structure. The branches proliferate toward the future and never toward the past because of the initial condition of kinematical independence.

In fact, the very "passing" of time has a realization in quantum theory only *via* the interpretation basis: In virtue of the deterministic dynamical evolution, any description of a quantum system at one instant is isomorphic to the description at any other instant. Thus it might seem that "nothing ever happens." We conventionally choose to describe the state at any time t in terms of eigenstates of "familiar" observables at the same time t, and thus obtain the appearance of time passing. But since any observable at one time is expressible in terms of observables at any other time, it would again seem that this appearance is illusory. Of course we now know that the proper choice of basis for the description of the system is not always a matter of arbitrary choice or convention. Here, we must use the interpretation basis. ("Familiar" in the above must be replaced by "diagonal in the interpretation basis.") The motion of the interpretation basis is the only thing in quantum theory that gives reality to the passing of time.

$S(t)$ is not an observable: After all, (73) and (74) are c-number equations. Nor should we expect the true physical entropy of the world to be an observable; for it measures no property of a single universe or branch, but rather the degree of branching of the state of the world as a whole. It is the evolution of the world as a whole which possesses a consistent forward arrow of time, even though a few ("maverick") branches will always behave differently from the majority. In the ensemble approximation this is mirrored by the fact that individual members of the ensemble always evolve reversibly whereas the ensemble as a whole may evolve irreversibly. The relationships between $S(t)$ and observable quantities such as coarse-grained entropies depend on the state and action functional of the world. For any particular quantum cosmological model these relationships could be elaborated, and the physical significance of the motion of $S(t)$ inferred.

Why are the interpretation basis states in the real world usually very close to the eigenstates of "classical" observables?"[3] This is closely related to the correspondence principle which I touched on in Section 7. This problem well deserves further study (perhaps along the lines of Daner *et al.*, 1962), but it is tempting to speculate as follows: *Whatever* the nature of the state and action functional of the real world had been, there would have been an interpretation basis. This would have singled out a preferred class of observables with respect to which classical laws of physics (i.e., laws assigning single objective values to each of these observables at every instant) would be approximately true in most universes.

[3]I am grateful to Prof. A. Shimony for raising this question.

ACKNOWLEDGMENTS

The research for this work was supported in part by NSF grant No. PHY 7826592. The author has benefited greatly both from Prof. B. S. DeWitt's profound understanding of quantum theory and from his constant encouragement. It is a pleasure to thank Prof. B. d'Espagnat, Prof. A. Shimony, and Prof. J. A. Wheeler for many useful discussions and for helpful criticism.

REFERENCES

Bell, J. S. (1964). *Physics*, **1**, 195.

Bohm, A. (1980). Decaying states in the Rigged Hilbert space formulation of quantum mechanics, *J. Math. Phys.* **21**.

Bohr, N., and Rosenfeld, L. (1933). *Kongelige Danske Videnskabernes Selskab Matematisk Fysiske Meddelelser*, **12**(8).

Daneri, A., Loinger, A., and Prosperi, G. M. (1962). *Nucl. Phys.* **33**, 297.

Davies, E. B. (1976). *Quantum Theory of Open Systems*, Academic Press, London.

d'Espagnat, B. (private communication).

d'Espagnat, B. (1976). *Conceptual Foundations of Quantum Mechanics*, second ed. W. A. Benjamin, Reading, Massachusetts.

DeWitt, B. S. (1968). The Everett-Wheeler interpretation of quantum mechanics, in *Battelle Rencontres, 1967 Lectures in Mathematics and Physics*, C. DeWitt and J. A. Wheeler, eds. *The Many-Worlds Interpretation of Quantum Mechanics.* New York.

DeWitt, B. S. (1970). *Physics Today*, **23**, 9.

DeWitt, B. S. (1973). The many-universes interpretation of quantum mechanics, in *The Many-Worlds Interpretation of Quantum Mechanics*, B. S. DeWitt and N. Graham, eds. Princeton University Press, Princeton, New Jersey.

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

Einstein, A., Podolski, B., and Rosen, N. (1935). *Phys. Rev.* **47**, 777.

Everett, H. (private communication).

Everett, H. (1957). *Rev. Mod. Phys.* **29**, 454.

Feynman, R. P., Leighton, R. B., and Sands, M. (1965). *Lectures on Physics*, Vol. 3. Addison Wesley, Reading, Massachusetts.

Gel'fand, I. M., and Shilov, G. E. (1968). *Generalized Functions.* Academic Press, New York.

George, C., Prigogine, I., and Rosenfeld, L. (1972). The macroscopic level of quantum mechanics, *Kongelige Dansk Videnskabernes Selskab Matematisk Fysiske Meddelelser*, **38**, 1.

London, F., and Bauer, E. (1939). *La Theorie de l'Observation en Mecanique Quantique.* Hermann & Cie., Paris; (English translation) *Quantum Theory of Measurement*, J. A. Wheeler and W. H. Zurek. Princeton University Press, Princeton, New Jersey (1982).

Misra, B., Prigogine, I., and Courbage, M. (1979). Lyspunov variable; Entropy and measurement in quantum mechanics, *Proceedings of the National Academy of Science*, **76**, 4768.

Popper, K. R. (1967). In *Quantum Theory and Reality*, M. Bunge, ed. Springer, New York.

Schmidt, E. (1907). *Math Annalen*, **63**, 433.

Schrödinger, E. (1935). *Proceedings of the Cambridge Philosophical Society*, **31**, 555.

Turing, A. M. (1950). *Mind*, **59**, 433.

von Neumann, J. (1930). *Fundamentos Matemáticos de la Mecánica Cuántica.* Institute Jorge Juan, Madrid.

von Neumann, J. (1932/1969). *Mathematische Grundlagen der Quantenmechanik.* Springer, Berlin. (Dover, New York, 1943).

von Neumann, J. (1946). *Les Fondements Mathématiques de la Mécanique Quantique.* Alcan, Paris.

von Neumann, J. (1955). *Mathematical Foundations of Quantum Mechanics.* Princeton University Press, Princeton, New Jersey.

von Neumann, J. (1964). *Matematičeskije Osnovi Kvantovoj Mekhaniki,* Nauka, Moscow.

Wheeler, J. A. (1957). *Rev. Mod. Phys.,* **29** (3), 463.

Wheeler, J. A. (1977). Genesis and observership, in *Foundational Problems in the Social Sciences,* Butts and Hintikka, eds., p. 3. D. Reidel Publishing Company, Dordrecht.

Wigner, E. (1961/1962). Remarks on the Mind-body Question, in *The Scientist Speculates,* I. J. Good, ed., Chap. VI. Heineman, London/Basic Books, New York.

Zeh, H. D. (1973). *Found. Phys.,* **3**, 1.

Zeh, H. D. (1980). Letter to J. A. Wheeler (unpublished).