Journal of Statistical Physics, Vol. 42, Nos. 3/4, 1986

Deterministic Quantum Evolution through Modification of the Hypotheses of Statistical Mechanics

L. S. Schulman[1,2]

Received April 1, 1985; revision received August 20, 1985

It is claimed that for all apparatus capable of performing macroscopic measurements of microscopic systems there exist special internal states for which deterministic quantum evolution alone yields a particular macroscopic outcome rather than a superposition of macroscopically distinct outcomes. We maintain that these special states are distributed uniformly (in a certain sense) among the set of all states. It is hypothesized that for all actually performed experiments the initial conditions lie among the special states. We postulate that in the absence of precise information on apparatus initial conditions one should give equal weight to those microstates that are consistent with the macroscopic state *and* are special in the sense used above. Evidence is presented for this postulate's recovering the usual quantum probabilities. This theory is fully deterministic, has no collapsing wave functions, and offers a resolution of the quantum measurement problem through a revision of the usual statistical mechanical handling of initial conditions. It requires a single wave function for the entire universe and an all encompassing conspiracy to arrange the right sort of special wave function for each experiment. In other words, an apparatus is in an appropriate microstate for the experiment that will actually happen even if an ostensibly random process is used to determine that experiment from among apparent alternatives. Although we do not provide physical or philosophical justification for our central hypothesis, some perspective is given by examining the notions implicit in the usual principles of thermodynamics.

KEY WORDS: Quantum measurements; thermodynamic arrow of time.

[1] Physics Department, Technion, Haifa, Israel (permanent address).
[2] Institutt for teoretisk fysikk, Universitetet i Trondheim, 7034 Trondheim NTH, Norway.

0022-4715/86/0200-0689$05.00/0

1. INTRODUCTION AND STATEMENT OF THE CENTRAL HYPOTHESIS

Quantum measurement theory deals with the reconciliation of the microscopic with the macroscopic, a task which in a different context is associated with statistical mechanics. The definiteness of measurement at the macroscopic level is commonly explained either by momentary suspension of the usual quantum evolution or by leaving the results of individual experiments undetermined by the theory, with the wave function bearing information about ensembles only.[1] In this paper we propose that the reconciliation can be achieved by a radical revision of our fundamental beliefs about *a priori* probabilities, and that this relieves us of any need to suspend quantum evolution as well as allows the use of the wave function to predict the results of individual experiments. The essential idea is that for a macroscopic measuring apparatus there exist microscopic states for which definite outcomes result under pure quantum evolution and that in actually performed experiments one or another of these special states occur as initial conditions.

Consider a Stern–Gerlach apparatus that measures the z component of the spin of silver atoms that pass through it. The apparatus is assumed to perform a complete measurement, that is, it consists not merely of a magnet with an inhomogeneous field but contains recording apparatus as well. For didactic purposes I will imagine the recording performed in an unusual but feasible way. The magnet sorts the beam into upper (say, for $s_z = +\hbar/2$) and lower beams and there are exit counters that respond to the presence of the Ag atom without rotating its spin. A pair of counters is placed so that an atom with appropriate spin triggers one or the other. (We assume beam intensity low enough to allow the recording of individual events.) The signal is sent to a typewriter that automatically prints the result of the experiment (e.g., "The particle triggered the upper counter and has spin $+\hbar/2$"). Moreover, if this is necessary, we include a physicist watching the typewriter as part of the apparatus and with each burst of printing the page is scrutinized. During the time the measurement is taking place both system and apparatus are isolated.

For a quantum mechanical description of the experiment, let φ_+ and φ_- be states of the silver atom which are (respectively $+\hbar/2$ and $-\hbar/2$) eigenstates of s_z, the z component of spin. Let ω be a state of the apparatus before the experiment and as such a point in a Hilbert space of enormous complexity. Suppose the initial state is $\varphi_+\omega$ (the product form is adopted

[1] In this one sentence summary I do not attempt to convey the subtleties of the many views of measurement theory that have emerged in the last 60 years. See Refs. 1–3.

for convenience only). If H is the total Hamiltonian and T the time the atom takes to pass through the apparatus then the final state is

$$\varphi'_+\omega'_+ = U(\varphi_+\omega) \qquad \text{with } U=\exp(-iHT) \tag{1.1}$$

where φ'_+ is a spin-up state of the atom (the prime takes note of changes in spatial coordinates, for example the fact that the atom now has a slight upward motion conferred by the magnetic field). ω'_+ is the final state of the apparatus and the "+" emphasizes that it has macroscopic properties (the printed page) that testify to the atom's passage through the upper exit channel. Similarly

$$\varphi'_-\omega'_- = U(\varphi_-\omega) \tag{1.2}$$

Consider an atom initially polarized along the $+x$ axis. With normalized φ's its wave function is $(\varphi_+ + \varphi_-)/\sqrt{2}$. By the superposition principle the experiment has the following effect on it:

$$\frac{1}{\sqrt{2}}\varphi'_+\omega'_+ + \frac{1}{\sqrt{2}}\varphi'_-\omega'_- = U\left[\frac{1}{\sqrt{2}}(\varphi_+ + \varphi_-)\,\omega\right] \tag{1.3}$$

The quandary of measurement theory is somehow to go from the coherent superposition in (1.3) to measurements that definitely give one of the two answers, + or −. For purposes of comparison we give a brief review of one conventional version of that justification. It is easiest to speak in terms of density matrices (see Ref. 2), which for the final state in (1.3) is

$$\rho = \begin{pmatrix} \omega'_+\varphi'_+\varphi'^{\dagger}_+\omega'^{\dagger}_+ & \omega'_+\varphi'_+\varphi'^{\dagger}_-\omega'^{\dagger}_- \\ \omega'_-\varphi'_-\varphi'^{\dagger}_+\omega'^{\dagger}_+ & \omega'_-\varphi'_-\varphi'^{\dagger}_-\omega'^{\dagger}_- \end{pmatrix} \tag{1.4}$$

Comparison of ρ with experiment is achieved by forming $\mathrm{Tr}(\rho A) = \langle A \rangle$ where A is an observable to be measured. In this trace the contribution of the off diagonal matrix elements of ρ has the form $\varphi'^{\dagger}_+\omega'^{\dagger}_+ A\omega'_-\varphi'_-$. It appears to be an empirical fact that for ω'_+ and ω'_- macroscopically different it is not possible to produce laboratory apparatus to measure any observable for which $\varphi'^{\dagger}_+\omega'^{\dagger}_+ A\omega'_-\varphi'_-$ would be nonzero.[2] Therefore ρ can be replaced by an effective $\hat{\rho}$ consisting only of the diagonal elements of ρ, and the pure state ($\mathrm{Tr}\,\rho = \mathrm{Tr}\,\rho^2 = 1$) has effectively become mixed ($\mathrm{Tr}\,\hat{\rho}^2 < 1$). And now, according to the statistical interpretation, one cannot for an individual experiment, make any further statements. Which nonzero entry on the diagonal of ρ actually occurs is not predicted by quantum mechanics.

[2] As mentioned in Ref. 2, the phenomena of superconductivity and superfluidity cloud the validity of this statement.

The first step in the theory to be presented in this paper is an assertion on the existence of certain special states of the apparatus. Although for the vast majority of apparatus states ω equations (1.1)–(1.3) are true to excellent approximation, I claim that there exist special apparatus states $\tilde{\omega}$ with the property

$$U\left[\frac{1}{\sqrt{2}}(\varphi_+ + \varphi_-)\,\tilde{\omega}\right] = \varphi'_+\,\tilde{\omega}'_+ \tag{1.5}$$

That is, the particle goes out the upper exit channel and its wave function has support there only. Similarly there are (other) $\tilde{\omega}$ for which the right-hand side of (1.5) is $\varphi'_-\tilde{\omega}'_-$. The special states $\tilde{\omega}$ are macroscopically indistinguishable from ordinary states of the apparatus but have built in many precise coherences that manage to produce (1.5) (representing a definite result for the measurement).

This assertion can be checked in principle within the context of ordinary quantum mechanics and in no way depends on interpretations. In this sense our theory is falsifiable: if there is an apparatus that makes definite measurements whose internal space is not rich enough to provide appropriate special states then the theory must fall. In later sections I shall argue for the existence of such states but for now let me remark that within statistical mechanics there are states that behave peculiarly but they are a tiny minority (of nonzero measure). There are states for which a raindrop collects on the ground and flies upward: just time reverse a fallen raindrop and the ground that received it and the air that picked up the faint sound of its fall. It turns out that the special states I need are harder to come by.

Assume then that special states for definite measurements exist. I now make the central hypothesis of the theory: In all actual experiments the initial state is special; the microscopic state of the apparatus is perfectly attuned to giving a definite result for the particular input wave function of the system to be measured. "Actual" means an experiment that is actually done, not contemplated, not merely calculated and not idealized, for example, by replacing the complex currents giving rise to the magnetic field by a fixed field function $\mathbf{B}(\mathbf{r})$.

Later in this paper I will try to make the foregoing hypothesis less unpalatable or at least pin down which of our fundamental intuitions it violates (it is the thermodynamic arrow of time that is seriously modified). But for now I want to convey the far reaching and perhaps disturbing consequences of our claim. It says that for each experiment the apparatus is ready for the system it will measure. If there is a beam then the final apparatus state after one measurement is a special state for the next incoming atom. Since for the spin polarized in the y direction [e.g.,

$(\varphi_+ + i\varphi_-)/\sqrt{2}]$ a *different* special apparatus state is needed, it claims the apparatus will be in the correct special state even if a y-polarized beam is suddenly fed into it, even if this sudden change is the result of an apparently whimsical decision of a human experimenter or random decay of a nucleus. Continuing this line of reasoning in what some may consider a *reductio ad absurdum*, I require that the entire universe be one tightly, coherently interconnected system, a single wave function. In particular each apparatus arranges itself in rare states so as to provide definite output for the atom coming its way. There can be no deviation from the plan and all time evolution is deterministic evolution under the Schrödinger equation (or Dirac or QCD, etc.). The subjective perception of being able to control an experiment or change it at will must be considered in the present theory to be illusory. There is no delayed choice experiment because there is no choice. This paper is not the place to discuss the issue of free will and determinism or to wonder why we do or should strive when all is determined. In the context of classical mechanics these issues already arose and have long been the subject of philosophical discourse.[4–6]

There is a sense in which this is a hidden variable theory, in that the particular internal state of the apparatus fixes the outcome of the measurement. But I would distinguish this from what is usually meant by hidden variables. I neither add degrees of freedom that are not already required nor is there any hidden pointer within the atom or anywhere else that says the initial superposition $(\varphi_+ + \varphi_-)/\sqrt{2}$ is secretly (say) spin-up. At the end of the experiment it will become spin-up but that will be by normal interaction with electrons and other components of the apparatus that are poised to provide just the right pushes and pulls. The idea that the apparatus could force definite measurements was considered and proved impossible by Komar[7] but he did not go so far as to suppose that for each experiment the initial conditions would be specifically tailored to a particular outcome.

Having required a grand wave function for the universe (and I believe the isotropy of background radiation suggests a causally linked early cosmology) and with it an all reaching conspiracy to provide definite results for measurements, the question arises as to why the usual regularities occur. Why, when a large number of seemingly identical experiments are considered, should this detailed conspiracy nevertheless give the usual ratios of different outcomes, namely, the probabilities given by ordinary quantum mechanics in which the details of the apparatus are ignored? For this we shall require an additional postulate, a variant of that used in statistical mechanics for similar purposes. The special states, although presumably rare among the set of all states, are nevertheless abundant. To predict the outcome of an experiment in which we know

firstly the macroscopic state of the apparatus and secondly that its microscopic state is *some* special state we give equal *a priori* probability to each macroscopically suitable *special* state and average over these alone. If, as we shall later argue, this recovers the usual quantum probabilities then it will turn out that probability in quantum mechanics is no different from probability in statistical mechanics, namely, a democratic treatment of ignorance of the *exact* initial state.

In the next section we argue for the existence of the "special" states required in the theory. In Section 3 we discuss the arrow of time and consider our central hypothesis in that context. Following that we deal with the issue of recovering the usual quantum probabilities. Section 4 uses the probability postulate to relate the probability density for locating a particle to the absolute value squared of its wave function. The demonstration is made in the context of a semiclassical approximation. The last section is a discussion and summary.

The radical hypothesis on the occurrence of special states as initial conditions in all actual physical processes naturally raises the question of why this should happen. Even if all my technical claims are substantiated there is still no indication of why special states should be favored or why this process should not suddenly stop and the world fade into a fuzzy superposition. As will be seen in Section 3, the postulating of special states is not unfamiliar and contemporary physics already includes such a notion. But the question remains and leads me to believe that what is presented in this paper is missing some important component.

Finally there is the issue of experimental test. Since I have nothing but quantum evolution it is hard to see how contrasting predictions could emerge from this theory and the standard one. What is needed is a situation where standard measurement theory makes predictions beyond what is contained in the quantum equations alone.

A brief description of the present theory has already been published.[8]

2. EXISTENCE OF SPECIAL STATES

A special state is one that gives a definite measurement. A definite measurement occurs when the final result of an experiment is a pure state that is not a superposition of two or more macroscopically distinct states. This is not so much a definition as a distillation whose vagueness is due to the term "macroscopic." This vagueness however, is no more limiting here than in other contexts (e.g., defining classical entropy) and we proceed under its shadow.

As in Section 1, we speak in terms of the Stern–Gerlach apparatus and

use the same notation. By the definition just given there are trivial special states. For example, if the atom initially has spin-up essentially any initial apparatus state will give a definite result.[3] The interesting case is when for most initial apparatus states quantum mechanics requires a superposition of macroscopically distinct components. It is then our claim that there are special (initial) apparatus states for which only one component is reached under ordinary quantum evolution.

We shall offer two sorts of argument. The first is qualitative and so general as to apply to all measurements. But our reasoning will be without force unless one accepts quantum evolution exclusively (no suspensions or variant dynamics for the period of wave function "collapse"). Moreover, it is a qualitative argument of the sort that lends itself to dispute. In the second line of reasoning we construct a model of a measurement device and actually exhibit the special states. The defect here will be that our gedankenapparatus is not a model of anything. It has properties suitable for a measuring apparatus but it would surely be preferable to have a real apparatus in mind.

Suppose then that a Stern–Gerlach apparatus is set to measure the z component of the atom's spin and an atom is sent through that has been prepared in the $+\hbar/2$ eigenstate of the x-component of spin. For some particular performing of the experiment suppose the atom passes through the upper counter (as described in Section 1) and is thereby determined to have s_z value $+\hbar/2$. Now even for the statistical interpretation of quantum mechanics it is always the case that *for each experiment the conservation laws are obeyed.* Therefore in the particular experiment mentioned above ($s_z \to +\hbar/2$) there is a transfer of linear and angular momentum from the apparatus to the atom. By putting the counters well downstream the spatial precision needed to distinguish the upper exit channel from the lower one can be made small and the significant transfer of the conserved quantities takes place at the magnet. In other words, the electrons that give rise to the field have picked up momentum and angular momentum, ultimately conveying them to other degrees of freedom of the isolated apparatus. So there was momentum transfer to (say) the electrons during the course of the bending that occurred in the particular experiment performed. Can that

[3] Recall however that Wigner, Yanase, and Araki have shown in papers appearing in Ref. 1 that it is impossible for the scheme always to work. For this reason and because the usual spin-up → spin-up result is good only in the adiabatic approximation, arbitrary initial conditions will in general lead to a small component of the wave function in the spin-down final state. Therefore even in this case it is not trivial to produce "special states" and it must be checked that the dimension of the subspace of states which go entirely to spin-up compared to the relatively small dimension going to spin-down has the ratio appropriate to the usual calculated wave function. (See Section 3 for discussion of probabilities and subspace dimension.)

transfer be described by quantum mechanics? It will not be an easy calculation but it falls in the domain of quantum mechanical evolution no less than does the momentum transfer calculation for Compton scattering. Of course the final passage through the counters will seriously dislocate the phase relation between the final atom state and its state before it triggered the counter. But that passage does not eliminate the acquisition of momentum by the magnet part of the apparatus nor does it preclude the quantum mechanical description of that acquisition.

We continue to work from the end to the beginning. The upward moving (in positive time) atom[4] and disturbed magnet are next followed back through their quantum evolution to their respective states just prior to the atom's entering the inhomogeneous field. The state of the atom is known: $(\varphi_+ + \varphi_-)/\sqrt{2}$, by preparation. The state of the apparatus, including magnet, at this early time is something, call it ω_0. But this ω_0 has the property that if we now go forward in time it takes the system from $(\varphi_+ + \varphi_-)/\sqrt{2}$ to φ'_+ alone. It is therefore a "special" state.

As indicated earlier, I would not be surprised if the well-honed interpretive resources of the statistical approach would find flaws in the foregoing so I expect I am not doing more than preaching to the converted. Moreover, the demonstration may fall through completely in the face of Everett's "relative state" interpretation[9] of quantum mechanics.

One consequence of the existence of special states for the Stern-Gerlach apparatus is that there are initial states for which an incoming spin-up state does not emerge as spin-up [e.g., use the $\tilde{\omega}$ of Eq. (1.5)]. It should be mentioned therefore that even a conventional treatment ignoring the microscopic coordinates of the magnet (and not invoking Wigner, Araki, and Yanase) yields an occasional flip since the usual conclusion is true only in the adiabatic approximation (see Ref. 2, Section 18). In fact were it not for this approximation a naive treatment of the transverse fields (B_x and B_y) necessarily accompanying the nonconstant B_z predicts that spin flips

[4] Here is where holders of the standard interpretation would find my argument unconvincing. In the usual description, immediately preceding the triggering of the counter the wave function had support in both channels and although the measurement ultimately revealed the particle to be in the upper channel it is meaningless to ascribe a particular channel to it before the measurement. But once again I appeal to conservation laws. Calling the vertical direction z, if the magnet conveys Δp_z to the atom and the detection process conveys δp_z (due to distinguishing between upper and lower channels), then so long as $|\delta p_z| \ll |\Delta p_z|$ knowing the momentum after the detection gives a range of possible momentum values before the detection that is sufficiently accurate to pin down the atom's previous channel. On the face of it, making the detector far from the magnet allows $|\delta p_z|$ to be arbitrarily small but having seen many subtle arguments of this sort I am prepared to admit the possibility that there is some unbeatable combination of uncertainties to preclude the p_z determination I require.

should often occur. This suggests that the place to start in looking for special states of this apparatus would be in the breakdown of the adiabatic approximation.

In the second approach to showing the existence of special states we consider a model of a measuring apparatus, show in what sense the model has properties suitable to such a designation, and display the special states. We call the apparatus a "detector" (or a small sensitive region within a detector) and call the system or object whose position is measured a "particle." But these names are a convenience and at the least our model suffers from being an oversimplification.

In the forthcoming discussion we shall talk about phonons, particles, atoms, etc. without the explicit quotation marks heretofore employed. This is mainly for convenience although it is hoped that the terminology is not completely irrelevant. The detector consists of large clusters of identical atoms grouped in many sensitive units of various shapes, sizes and orientations, like the grains in a film. The atoms are in a metastable level. With the passage of the particle they are strongly coupled to phonons in the environment and decay. These first-level phonons are in turn coupled to other modes (second-level phonons) and their energy is spread among them. The particle to be detected is an atom in its excited state. That is, we do not detect the atom, which is always assumed to have gone through the detector, but only whether, during its passage, it was in some particular excited state of interest. The traversing atom will therefore be treated as a two level system, its ground state the vacuum and $\varphi^\dagger$ the operator that raises it to the excited state of interest. The only role for its center-of-mass coordinates is to carry it through the detector so that the time interval during which a grain (an elementary subunit of the detector—a sensitive cluster) is exposed to the particle depends on its velocity and the size, shape, and orientation of the grain. Conceivably the center-of-mass coordinate could be deflected by detection, as in the Stern–Gerlach experiment, but we will ignore this and assume that the macroscopic measurement message will be carried by the second level of phonons. We also idealize the atoms of the detector as two-level systems and their associated creation operators are labeled $a_k^\dagger$, $k=1,\ldots, M$. The first-level phonons in the grain have creation operators $b_j^\dagger$, $j=1,\ldots, N$. These interact strongly with the atoms in the presence of the particle. For the creation operators of the other phonon modes, into which the b modes make their leisurely decay, we use the notation $c^\dagger$, although we shall not actually make any calculation involving them. The free Hamiltonian, when the particle is far from the detectors, will be taken to be of the form

$$H_0 = \sum_{k=1}^{M} E a_k^\dagger a_k + \sum_{j=1}^{N} \omega_j b_j^\dagger b_j \tag{2.1}$$

where considerable simplifications have already been made: The free Hamiltonian of the c's is ignored; their energies are well below E. The b–c coupling is ignored as it will be much weaker and play no role on the time scale for detection. The range of ω_j is assumed to be a narrow band around E, $[-\omega_0+E, \omega_0+E]$. We are omitting the energy difference of the states of the particle since we assume its effect already incorporated in the a and b energies. Within a particular cluster we assume that on the average one detector atom is excited initially and that all the atoms couple to all the b phonon modes in the same way. In particular the coupling is taken to be

$$H_1 = \frac{1}{\sqrt{N}} g \sum_{k=1}^{M} \sum_{j=1}^{N} b_j^\dagger a_k \varphi + \text{adjoint} \tag{2.2}$$

The factor $1/\sqrt{N}$ provides the correct scaling to allow a transition to a continuum for the b levels. We have in mind that $1 \ll M \ll N$.

If T is the time for the particle to traverse a particular cluster, the time evolution operator of interest is $U = \exp[-i(H_0+H_1)\,T]$. For an initial state $\theta = \varphi^\dagger a_k^\dagger |0\rangle$ the final state will be $U\theta$, where we do not indicate evolution under H_0 for t outside the interval $[0, T]$ since it does not affect the measurement. As mentioned above, b–c coupling is negligible during $[0, T]$ but its subsequent effect is that the b levels have a slow, exponential, irreversible decay[10] into the c levels. For sufficiently large energy difference between b and c levels there can be considerable amplification and a macroscopic event recorded.

Define the span of the vectors $a_k^\dagger \varphi^\dagger |0\rangle$, $k = 1,..., M$ to be L_a and that of $b_j^\dagger |0\rangle$, $j = 1,..., N$ to be L_b. Call the associated projection operators P_a and P_b, respectively.

In the conventional interpretation of quantum mechanics, the probability that a grain detects the particle will be $\|P_b U\theta\|^2$, that is, the norm squared of the amplitude characterizing the decay of the detector atoms. The probability of traversal without detection is the norm squared of the amplitude for remaining in L_a, namely $\|P_a U\theta\|^2$. Our goal in developing this model is to show that there exist special initial conditions $\tilde{\theta}$ of the model, consistent with the macroscopic state, for which it is transmitted completely, i.e., $P_b U\tilde{\theta} = 0$, and (other) special states $\tilde{\theta}$ for which it is absorbed completely, $P_a U\tilde{\theta} = 0$. The macroscopic state in this case is characterized by the fact that there is one detector atom per grain initially excited, that is, $\sum_{k=1}^{M} a_k^\dagger a_k$ has expectation value 1, and that the b's are not excited, that is, $\sum_{j=1}^{N} b_j^\dagger b_j$ is zero. It is likely that to obtain special states for some kinds of apparatus finite temperature effects will be needed, which for microscopic apparatus states translates into allowing some initial excitation in the b's, deviations in the distribution of excitation among the a's, (in dif-

ferent clusters) as well as a tolerance of imprecision in the final state. We shall not need such effects in the present model (except possibly the last) but that may be due to its artificiality and for realistic systems the superficially random thermal fluctuations may be important in the formation of special states, as they would be for our atypical raindrop mentioned in the Introduction.

The model presented is rather simple to analyze and this is done in Appendix A. The high degree of degeneracy (M) arising from the weakly coupled large number of detector atoms in the grain allows us to produce a class of perfectly transmitting states, namely, all θ of the form

$$\tilde{\theta}_T = \sum_{k=1}^{M} \varphi^\dagger \gamma_k a_k^\dagger |0\rangle, \qquad \sum_{k=1}^{M} \gamma_k = 0, \qquad \sum_{k=1}^{M} |\gamma_k|^2 = 1 \tag{2.3}$$

The span of such states forms an $(M-1)$-dimensional space in which each point is an eigenstate of H_0 and of $\sum_{k=1}^{M} a_k^\dagger a_k$ with eigenvalue 1, and is also an eigenstate of $H_0 + H_1$ so that the evolution U does no more than multiply $\tilde{\theta}_T$ by a phase factor; moreover, $P_a U \tilde{\theta}_T = U \tilde{\theta}_T$.

The overresponsive state for which $U\tilde{\theta}$ lies entirely in L_b can be gotten from the remaining dimension of L_a, namely,

$$\tilde{\theta}_A = \varphi^\dagger M^{-1/2} \sum_{k=1}^{M} a_k^\dagger |0\rangle \tag{2.4}$$

may be totally absorbed. However, as shown in Appendix A it is not totally absorbed at all times but only at T such that

$$gT\sqrt{M} = (n + \tfrac{1}{2})\,\pi, \qquad n = 0, 1, \ldots \tag{2.5}$$

And even so the absorption is incomplete by a term of order ω_0^2/g^2M. Therefore in a given experiment and in a given piece of the macroscopic detector not all grains will be suitable for total absorption. If the detector is rich enough in the variety of its clusters to have one for which the traversal time (of that particular cluster) and degeneracy are suitably related then it will in fact have the special state needed for absorption. As to the $O(\omega_0^2/g^2M)$ discrepancy, it is very small (I expect large values of M and rather strong coupling during the period of interaction) and I do not know whether in principle it must be exactly zero. Perhaps this is where thermal fluctuations play a role.

These then, subject to the deficiencies indicated, are the special states needed for total transmission or detection of the particle.

Note finally that we have not considered initial states with components in the space Span$(a_k^\dagger |0\rangle)$, that is, with no initial excitation of the atom. In our model there is no possibility of absorbing these vectors com-

pletely since no mechanism by which this state communicates with L_a and L_b is given. This would be the analog of neglecting deviations from the adiabatic approximation (or other similar features) in the Stern–Gerlach experiment.

3. HYPOTHESES ABOUT MICROSTATES

The central hypothesis of the theory presented here is that for each actual experiment the initial microscopic state is "special," that is, is one of those that gives a definite macroscopic result. Although I can offer neither justification nor rationale for this assumption, I do wish to give perspective. What will now be shown is that in ordinary statistical mechanics we are already familiar with the idea that given a macrostate only certain particular microstates can appear. Such a restriction on the final state in an isolated experiment is standard. The radical step of the present theory is to place restrictions on initial conditions as well. The radical step is thus interpreted as a departure from the thermodynamic arrow of time or alternatively as a change in our rules for assigning *a priori* probabilities.

First we present a particular way of defining the thermodynamic arrow of time.[11] Suppose that at 2 p.m. there is a cup of water with an ice cube floating in it. The system is left isolated for an hour and we wish to predict the state at 3 p.m. By the principles of statistical mechanics we proceed as follows: Consider all microstates consistent with the macroscopic description (water temperature, etc.) at 2 p.m. Let each of these evolve by appropriate microscopic equations of motion and then average over the resulting 3 p.m. microstates. The result is less ice and colder water. A small minority of microstates do not conform but their effect is lost in the averaging. (For a model on which these calculations can be implemented see Kac.[12]) This is an example of the fundamental principle giving equal *a priori* probability to all microstates consistent with a given macrostate.

Suppose further that the system was isolated for an hour prior to 2 p.m. and we inquire as to its 1 p.m. state. Giving equal *a priori* probability to all microstates consistent with the 2 p.m. macrostates would give less ice and colder water at 1 p.m. as well, by the time symmetry[5] of the evolution. But all experience indicates that there is more ice and warmer water at 1 p.m., so there clearly must be a different rule for handling probabilities when deducing the past (retrodicting). It is this: Consider hypothetical macrostates at 1 p.m. and evolve them forward using exact

[5] Full-time symmetry is not needed to reach this conclusion, as for example in Ref. 12 where the inverse process differs from the forward process but time reversal paradoxes still hold. Similarly, I do not expect T violation in weak interactions to invalidate my conclusions.

dynamics and equal probabilities for all microstates consistent with the hypothetical 1 p.m. macrostate. One p.m. macrostates that yield the 2 p.m. macrostate in this way are possible 1 p.m. states. This asymmetry in the way of handling prediction and retrodiction is one way of phrasing the thermodynamic arrow of time.

Coming back to an observed 2 p.m. macrostate, what can be said of its microstates? The answer depends on whether that state is considered an initial or final condition. In particular, as a final state only a small (but finite measure) and precisely determined subset of the microstates is acceptable. That precision is beyond all conceivable experimental art. In other words, if you come across a floating ice cube that was isolated for some period prior to your observation, to the extent that you try to imagine its earlier condition, you must ascribe to Nature a conspiracy[6] to have it in those special microstates that lead backwards to reasonable precursors, larger ice cubes, warmer water.

Our central hypothesis, which places severe restrictions on possible initial conditions, is therefore not so far fetched in that it limits microstates from among many plausible candidates but in that it does this for *initial* conditions. In this sense our central hypothesis is a denial of the apparent arrow of time.

The next question we take up has to do with recovering the ordinary. Why is it, if there is an all encompassing conspiracy to have every initial condition be special and rare, that things do not look that way? If every precursor is uniquely and improbably selected, how do we recover the usual regularities? Or, to put the question more concretely: Why if an experiment is performed many times are the results given by the usual probabilities of quantum mechanics? Why, for example, if an s_x-polarized beam is sent into the Stern–Gerlach apparatus described above will roughly half the events give s_z-up and half s_z-down?

For this we require an additional postulate, a modification of the usual probability rule of statistical mechanics. Consider the region of phase space W associated with some macrostate of the apparatus.[7] It has a con-

[6] The "conspiracy" is similar to that needed for our theory in that far-flung events in the universe must be coordinated and the particular microstate takes into account distant occurrences as well. For example, suppose our nominally isolated glass of water and ice was struck by a cosmic ray at 1:30 p.m. Then the 2 p.m. microstate used to get a bigger ice cube at 1 p.m. must include this event for if it did not the rapid divergence and instability of microscopic evolution would cause the ice cube to shrink as we got to earlier and earlier times before 1:30.

[7] For ease of presentation the discussion is phrased in classical terms but the implementation, as for example in the "detector" model to be given below, is quantum mechanical. The "convoluted subsets" of the classical description are subspaces of the apparatus Hilbert space in the quantum treatment.

voluted subset W' that corresponds to the special states. Ordinarily in making macroscopic predictions one averages over W. We postulate that macroscopic predictions are gotten by averaging over W' alone. That is, to make predictions from initial macroscopic conditions give equal weight to all microstates that are both "special" and consistent with the macrostate. ("Special" here has the meaning defined earlier in this paper.) We claim that with this postulate the outcomes of an oft-repeated experiment are given by the usual squared amplitudes of the system to be measured. This is a strong assertion since it requires the relative weight of apparatus states capable of achieving a particular "definite" result to be proportional to squared amplitudes in the system to be measured.

As for our earlier statements on the existence of "special" states we will try to justify our claim in two ways—the first general but incomplete and the second based on our "detector" model and thereby weakened by its limitations. Before beginning this justification however we wish to consider some implications of the postulate.

Quantum probability, by this postulate, is no different from classical probability; both arise from ignorance of exact initial conditions. Just as for classical dice, if in quantum mechanics we know the exact initial point in phase space we would know the outcome; in the absence of such knowledge the probability of various outcomes follows the relative weight of acceptable initial states (acceptability for quantum systems includes the property of giving "definite" outcomes).

Next we address one of the less precisely worded versions of the question posed earlier. Why if only a tiny subset of phase space is allowed, does everything look normal? For this problem I will review an example studied some years ago[13] in which initial conditions are an intricately interwoven subset of a macroscopic subset of phase space and in which for all purposes not related to the specific definitions of that interweaving the system behaves normally—i.e., as if averaging were being done over the entire macroscopic subset.

The system studied was the well-known[14] automorphism Φ of the unit square $\Phi(x) \equiv x + y$, $\Phi(y) \equiv x + 2y$, both assignments modulo 1. The fact that this transformation is mixing makes it easy to see equilibration, for example in the following way. Make a coarse graining, that is, divide the square into $M \equiv N^2$ grains where grain (i, j) is the subset $G_{ij} \equiv \{(x, y) \mid i-1 \leqslant Nx < i, j-1 \leqslant N_y < j\}$, $i, j = 1, \ldots, N$. Specifying a grain will be taken to represent the providing of the best macroscopic information available. If we watch the evolution of the points in a single grain they will rapidly spread throughout the unit square. This can be made quantitative by defining the entropy $S(V)$ of a set V with respect to the coarse graining.

Let[8]

$$S(V) \equiv -\sum_{i,j}\left[\frac{\mu(V \cap G_{ij})}{\mu(V)}\right]\log\left[\frac{\mu(V \cap G_{ij})}{\mu(V)}\right] \tag{3.1}$$

where μ is Lebesgue measure on $\mathbb{R}^2$. Define the tth iterate of Φ to be Φ_t, i.e., $\Phi_0 = 1$, $\Phi_1 = \Phi$, $\Phi_t = \Phi \circ \Phi_{t-1}$. Now take V_t to be $\Phi_t(G_{11})$. Clearly $S(V_0) = 0$. Moreover, it follows from the mixing property that $\mu(V_t \cap G_{ij})$ tends to $1/M$ as $t \to \infty$ so that $S(V_t)$ tends to $\log M$ corresponding to complete lack of information about the system's location. (In making analogies to classical mechanics the square should be thought of as phase space.)

In Ref. 13 we studied the effect of giving information about the system at two different times. Our purpose was to give symmetric boundary conditions in time (rather than initial conditions) so that other possible sources of asymmetry (such as expansion of the universe) could be considered. The procedure was to start the system in one grain, say G_{11}, and insist that at some fixed time $T(>0)$ it reach a particular (possibly other) grain, say G_{24}. This was implemented computationally[15] by starting a point in G_{11}, taking T time steps, and discarding all points that did not end in G_{24}. Analytically the set of acceptable initial conditions was $G_{11} \cap \Phi_{-T}(G_{24})$. Call this set V' and its forward images in time V'_t, $t = 1, 2, \ldots$.

We have recalled this example to mention one property of V'_t: It equilibrates in the same way, same shape, same time constants, as V_t. That is, for small t (and large T) the function $S(V'_t)$ is indistinguishable from $S(V_t)$: both make their way from 0 to a fluctuating behavior around $\log M$. Eventually, as $t \to T$, $S(V'_t)$ will begin to decrease as V'_t collects itself into G_{24}. But the detailed, sensitive information needed to accomplish this was not in evidence for small t and the special subset V' could as well have been all of V.

Now suppose that instead of averaging over all of V we averaged over a set of points randomly selected from V. For an independent selection process all expectation values should converge to the averages over V as the number of points goes to infinity. We claim that our obtaining ordinary behavior for $S(V'_t)$ for small t follows from the pseudorandom distribution of the points of V' among those of V. By "pseudorandom" we mean that even though V' is deterministically selected by time T properties, correlations between time T and time 0 are greatly reduced due to the mixing property of the transformation (for large T). Therefore for sets and observables defined in terms of time T macroscopics the deviation from true randomness is nil.

Our "special" states for the definiteness of quantum measurements are analogous to those just defined in that they are characterized by the

[8] This is not the Kolmogorov entropy defined in Ref. 14.

demand that at some (far) future time the system have well-defined macroscopic properties. We will use this analogy at two levels. The first level is simply to get one used to the idea that a system can have hidden internal correlations but that if one studies observables not specifically attuned to these correlations they will be invisible.

The second level brings us back to a demonstration promised earlier, namely, showing that averaging over the "special" states of the apparatus restores the probabilities obtained in the standard interpretation of quantum mechanics. Now we declare at the outset that the first set of arguments we shall give is incomplete. But a weak proof does not imply a false theorem and we expect there to be salvageable parts to our demonstration.

The discussion will be in terms of the Stern–Gerlach apparatus. To say that the result of a measurement was "spin-up" means both that the typewriter printed the appropriate statement and that the atom's wave function was more or less entirely located at the upper exit channel. Therefore if $\chi_+(\chi_-)$ is the characteristic function (as a function of the atom's spatial coordinates) of the region around the upper (lower) exit channel, then at the end of the experiment the (scalar) quantity $\varphi'^{\dagger}_{+}\omega'^{\dagger}_{+}\chi_+\omega'_+\varphi'_+$ has the value 1. Now suppose the initial atomic state is some $\theta = \alpha_+\varphi_+ + \alpha_-\varphi_-$ with $|\alpha_+|^2 + |\alpha_-|^2 = 1$. Then the special states are those $\tilde{\omega}$ for which $U(\tilde{\omega}\theta)$ is either $\varphi'_+\tilde{\omega}'_+$ or $\varphi'_-\tilde{\omega}'_-$. It is our goal to show that the relative abundance of the respective kinds of $\tilde{\omega}$ is proportional to $|\alpha_+|^2$ and $|\alpha_-|^2$. Using the notation introduced above, this means we must show that

$$J_\pm \equiv C^{-1}\sum_{\tilde{\omega}} \langle U(\tilde{\omega}\theta)|\, \chi_\pm\, |U(\tilde{\omega}\theta)\rangle = |\alpha_\pm|^2 \tag{3.2}$$

where the sum is over all special states and the normalization C is the same in both equations. Define a function $\tilde{\rho}(\omega)$, $\omega \in \Omega$ to be $1/C$ for special states (for the particular experimental setup, i.e., the particular $\alpha_\pm$), zero otherwise. (We ignore technical problems such as infinities in C.) Then the sum in (3.2) can be taken over all states $\omega \in \Omega$

$$J_\pm = \sum_{\omega} \tilde{\rho}(\omega)\langle U(\omega\theta)|\, \chi_\pm\, |U(\omega\theta)\rangle \tag{3.3}$$

The function $\tilde{\rho}(\omega)$ is extremely complicated, analogous to the (characteristic function of the) set V' discussed above. Should it happen that with respect to the expectation value of interest it has a pseudorandom character then $\tilde{\rho}(\omega)$ can be replaced by $\rho(\omega)$, the function on Ω that picks out those microstates of Ω consistent with the given macroscopic conditions. This is just the usual density matrix. Replace then $\tilde{\rho}$ in (3.3) by ρ and consider $U(\omega\theta)$. This is now a standard calculation and for the vast majority

of ω it is simply $\alpha_+\varphi'_+\omega'_+ + \alpha_-\varphi'_-\omega'_-$. Inserting this in (3.3) with ρ replacing $\tilde{\rho}$ yields

$$J_\pm = \sum_\omega \rho(\omega)[\alpha_+\varphi'_+\omega'_+ + \alpha_-\varphi'_-\omega'_-]^\dagger \chi_\pm [\alpha_+\varphi'_+\omega'_+ + \alpha_-\varphi'_-\omega'_-] \qquad (3.4)$$

The cross terms $\omega'^\dagger_\pm \omega'_\mp$ drop out by orthogonality of the macroscopic states and $\varphi'^\dagger_\mp \chi_\pm \varphi'_\mp = 0$ so that we have for $J_\pm$ the values $|\alpha_\pm|^2$, as desired.

There are two major lacunae in this argument. First, since $\varphi'^\dagger_+ \omega'^\dagger_+ \chi_+ \omega'_+ \varphi'_+$ is 1 (for any $\tilde{\omega}$) why could we not have chosen [in Eq. (3.2)] its 15th power rather than its first? This would have given a different answer in (3.4). Second, assumptions of pseudorandomness are appropriate when the conditioning that defines the pseudorandom subset is unrelated to the quantity being measured. In our case the conditioning is related to a distinguished time, the very time at which the expectation is evaluated. True, we are not conditioning on a particular outcome of the experiment, only that it have an outcome, but this still seems a serious liability.

Concerning the first lacuna, we shall later present evidence that when a wave function has some spread in space the probability of seizing the particle in a particular region is proportional to the absolute value squared of its wave function, locally. The evidence will be in the context of a scattering problem. Concerning the pseudorandomness we do not know how to improve our argument.

We next turn to the detector model to see how the relative availability of special states for various outcomes determines probabilities.

Suppose it is found that with a certain beam velocity, in a large detector consisting of many subunits of the sort described above ("grains"), the probability per grain of absorption is p_A, a rather small number. Let M be the closest integer to $1/p_A$ and consider clusters of size M [same "M" as in Eq. (2.1)] for which the passage time of atoms in the beam through the cluster satisfies $T = \pi/2g\sqrt{M}$. If there are no such clusters this apparatus will not serve as a measuring device. (In fact this is a general claim we make: the usual demands of amplification and practical irreversibility are not the only requirements for real measurement devices. In addition sufficient dynamical richness is needed to produce special states.) These particular clusters provide the desired special states. Consider first a particle that is detected. For this the internal state of one of these particular clusters is taken to be $\tilde{\theta}_A = \varphi^\dagger M^{-1/2} \sum a_k^\dagger |0\rangle$ of Eq. (2.4), giving a single dimension of the cluster Hilbert space for complete detection, as discussed in Section 2. All other clusters are in their totally transmitting internal states

(which require no tailoring of the passage time). For perfect transmission through the entire apparatus, i.e., no detection, our particular cluster is in the $(M-1)$-dimensional subspace of the cluster that provides perfect transmission [cf. Eq. (2.3)] and all other clusters are perfectly transmitting as well. The relative number of dimensions for detection compared to nondetection is therefore $1/(M-1)$ or $(1/M)/[(M-1)/M]$, which by the choice of M is $p_A/(1-p_A)$. It remains only to justify the identification of the counting of dimensions in Hilbert space with the assignment of equal *a priori* probabilities. But by the correspondence principle phase space volumes with characteristic dimension $\hbar$ to the appropriate power become individual levels for the quantum system, that is to say, dimensions in the Hilbert space. Thus our axiom for the assignment of *a priori* probabilities in the apparatus Hilbert space (count dimensions) is consistent with the classical assignment. Of course this is the same correspondence that justifies the trace operation (which gives equal weight to each dimension) in quantum calculations as the replacement for the phase space integral $\int d^N p\, d^N q$.

Finally, although we make no case for our model as an accurate description of any particular physical measuring device we wish to point out at least one reasonable property it possesses and respond to one possible criticism. Since the response of a single cluster is extremely small ($p_A \ll 1$) the full detector must have many such grains including many suitable for producing special states. Each such grain acts as a single coherent object independent of the others so that the probability of nondetection in two such clusters is the product of the probabilities for nondetection in each. This gives the detector the reasonable physical property of exponential dropoff of the probability for nondetection with increasing (macroscopic) length of detector material. (This is to be contrasted with increasing the size of individual grains for which such a property will not obtain.)

Next we respond to the following possible criticism: Consider an ordinary calculation of total absorption probability for this detector. There is some distribution of cluster sizes, shapes, perhaps also of coupling constant g. In each size L cluster one can perform a calculation of the sort given at the end of Appendix A. There will be probability $(L-1)/L$ of transmission because of degeneracy but now the Lth dimension is also partly transmitting because the passage time is not taking any special value (i.e., $T \neq \pi/2g\sqrt{L}$, necessarily). Thus if one goes to calculate overall transmission of a large detector, including features like the distribution of L's and time, why should it turn out that the transmission probability per grain be $1/M$? The fallacy of this question is that it reverses the logic of the selection of the M clusters. First we study the detector and calculate the absorption per grain to be p_A (also g is fixed by microscopic con-

siderations). After this we look for grains of order $[1/p_A]$. Thus we make no *a priori* assumptions on cluster size distribution except that there should be *some* grains of size M (but not that, e.g., the distribution peak there).

4. $|\psi(x)|^2$ AS POSITION PROBABILITY DENSITY

Our goal in this section is limited. It is not to show that we can produce special states for an arbitrary scattering experiment. Rather we wish to demonstrate that under certain particular conditions the absolute value squared of the wave function is the probability density for localization. In particular, and with the same disclaimers as to generality, we show that the relative dimension of special states that can localize a particle in a region around a point x is proportional to $|\psi(x)|^2$.

Consider a particle emitted by an appropriate device into a small spatial region around x_0. In the usual approach to this problem no features of the emitter would be called into play and the initial wave function would be some ψ_0 with a spread given by $\Delta_0 = [\int (y-x_0)^2 |\psi_0(y)|^2 \, dy]^{1/2}$. Upon emission the particle is subjected to a very slowly varying potential V. At some later time T an attempt is made to detect the particle in a neighborhood dx of a position x. The probability of success is $|\psi(x, T)|^2 \, dx$. $\psi(x, T)$ is gotten from ψ_0 using the propagator $G(x, T; y)$ and the circumstances outlined justify the use of the semiclassical approximation for G. If we further assume that for the given x_0, x, T, and V there are no turning points, ψ is given by [Ref. 16, Eq. (13.10)]

$$\psi(x, T) = \int dy \left(\frac{i}{2\pi\hbar} \frac{\partial^2 S}{\partial x \partial y} \right)^{1/2} e^{(i/\hbar)S(x,T;y)} \psi_0(y) \tag{4.1}$$

where $S = S(x, T; y)$ is the classical action, the integral of the Lagrangian along the classical path (solution of Euler–Lagrange equations) from y to x in time T. (To avoid confusion: S is here a function only of x, T, and y. It is the value of the functional $S[x(\cdot)]$ along the classical path determined by x, T and y.) If Δ_0 is small on all pertinent scales then we can simply approximate $\psi(x, T)$ by

$$\psi(x, T) = \left(\frac{i}{2\pi\hbar} \frac{\partial^2 S}{\partial x \partial x_0} \right)^{1/2} e^{(i/\hbar)S(x,T;x_0)} \tag{4.2}$$

(This formula breaks down, for example at large x, since ψ remains square

integrable.) Denoting the probability of finding the particle in the region dx around x by $P(x, dx)$ and taking note of the reality of S we have

$$P(x, dx) = \frac{1}{2\pi\hbar} \left| \frac{\partial^2 S}{\partial x \partial x_0} \right| dx \tag{4.3}$$

Now we ask whether there are special states into which the emitter can deliver the particle so that it finds itself at time T more or less entirely localized at x. Again we remind the reader that we do not now attempt to display "special" states for all scattering experiments, but are tailoring this problem to allow an interpretation of the wave function under at least some, perhaps unusual, circumstances. We assume then that for most of its internal states the apparatus produces a wave function ψ_0 which is far from minimal uncertainty and that although $\hbar/\Delta_0$ is large the spread in momenta is generally even larger than that. As special states we will take minimal uncertainty wave packets, thereby implicitly assuming that for rare initial conditions the emitter can produce them. The minimum uncertainty packets can be taken to be coherent states of an oscillator Hamiltonian whose ground state spatial uncertainty is Δ_0 and the states are labeled $|x_0, p_0\rangle$, the real and imaginary parts (modulo $\sqrt{2}$'s, etc.) of the usual complex label. Under Schrödinger evolution in the potential V, and again invoking the smoothness of V, these packets will follow the classical trajectory of a particle with initial momentum p_0, starting from x_0 (the oscillator Hamiltonian only defines states, not dynamics). If the original packet ψ_0 was badly nonminimal then our minimal packets need not spread significantly (both statements made with respect to the particular time T). If the detector at x has resolution poorer than the spread packet then these packets may serve as special states.

Next we count the relative dimension of the spaces of packets that arrive at various x, so as to compare to (4.3). Coherent states are not convenient for dimension counting since they are overcomplete, but as is well known an appropriately spaced regular lattice in the plane of complex labels does provide a countable complete basis set (there is one state too many but we ignore this). Since this label plane is essentially phase space for the particle we are considering, we can take the dimension of a collection of states $|x_0, p_0\rangle$ to be proportional to the classical phase space they occupy. This converts our problem into a simple question of classical mechanics: What range of initial momenta $[p_0, p_0 + dp_0]$ (with initial position x_0) brings the system into $[x, x + dx]$ at time T? In fact all we want is dp_0/dx, the relative number for each x. But here the role of the classical action as the solution of the Hamilton–Jacobi equation comes into play. p_0 is given by $-\partial S/\partial x_0$ (see Ref. 17 or 16). Since our interest is in the

change allowed this quantity for a specific variation in x we can in fact conclude that the relative probability for reaching various x is $|\partial^2 S/\partial x \partial x_0|$. That is,

$$\begin{aligned} &\text{Dimension of special states arriving in } [x, x+dx] \\ &\propto \text{phase space volume of initial momenta that} \\ &\quad \text{reach } [x, x+dx] \\ &\propto \left| \frac{\partial^2 S(x, T; x_0)}{\partial x \partial x_0} \right| dx \end{aligned} \tag{4.4}$$

This is proportional to $P(x, dx)$ of Eq. (4.3), justifying the use of the wave function for position probability density, at least under these special circumstances.

Having argued in what we hope has been a persuasive manner, we next discuss weaknesses and implicit assumptions. The dimensions we counted were in the particle Hilbert space. In principle we should count the dimensions of the emitter apparatus required for the production of various momenta. Since the emitter has not been described at all we can only offer the following justification. The high degree of localization achieved implies that momenta of interest here are small on the scale of energies involved in the localized emission process. Therefore if the required coherent states can be produced at all it should not much matter how much momentum must be imparted since so much is available. This gives equal emitter internal state space size for each p.

Next we point out that the special states we have presented must of necessity be flawed and some of our glib semiclassical approximations break down under circumstances only slightly different from those described above. For suppose it is possible that a second classical path can reach x from x_0. This can happen if V rises sufficiently sharply and T is long enough, circumstances that need not invalidate the usual semiclassical approach. The propagator will be a sum of two terms of the form $(S^\alpha_{xy})^{1/2} \exp(iS^\alpha(x, T; y)/\hbar)$ $(\alpha = 1, 2)$ with S^α the action along the respective classical path. Again for ψ_0 sharp enough the dominant behavior in the probability function will come from the interference term $\cos[(S^1 - S^2)/\hbar]$ with y evaluated at x_0. Superficially this looks like disaster for the proposed postulate that probabilities reflect dimensions of suitable special states. This is because each action S^α, $\alpha = 1, 2$, has a corresponding initial momentum $p^\alpha = -\partial S^\alpha/\partial x_0$ and the packets $|x_0, p_0\rangle$ can independently be sent to x. Thus the dimensions should add and no interference occur. Nor is there rescue from the claim that superpositions of these two special states would interfere, because they do not. This follows from the near orthogonality of the $|x_0, p^\alpha\rangle$ (the lattice coherent state basis is complete

but not orthonormal, but the overlap drops off sharply with the separation of the p's) and the unitarity of the evolution U. Thus

$$\langle U(x_0, p^1) | U(x_0, p^2) \rangle = \langle x_0 p^1 | x_0 p^2 \rangle \ll 1 \tag{4.5}$$

and despite the presence of both packets in a small neighborhood of x they remain nearly orthogonal.

But this last observation shows that if all my semiclassical assumptions were correct in the two-path case a standard calculation would show no interference either. For one could decompose the packet ψ_0 into a superposition of momentum states which in some approximation reflects a decomposition into coherent states of reasonably well-defined momentum. In evaluating the probability of detection we compute a matrix element $\langle \psi_1 | U | \psi_0 \rangle$ with ψ_1 a form determined by the detector. In this inner product—if our earlier semiclassical considerations are valid—all terms but those corresponding to the momenta p^1 and p^2 will drop out and because of the same orthogonality properties mentioned above, no interference will occur. That is, in the standard quantum calculation the interference term comes precisely from the overlap integral $\langle Up^2 | Up^1 \rangle$ and since this is presumed small here too there will be no interference. What is happening is that in order for the original argument to go through a badly nonminimal wave packet was needed and if such a sloppy emitter were used in a situation where reflection occurred the interference pattern would be wiped out.

The conclusion then is that the coherent states are not candidates for special states under all circumstances but if some circumstances do in fact allow their use, the square of the absolute value of the wave function emerges as a position probability density.

Finally, as alluded to earlier in this section, scattering theory in general presents a challenge, since from the moment of emission to that of detection a pair of scattered particles may not interact with any macroscopic apparatus and yet the general theoretical framework seems to require that the scattered wave function organize itself entirely into the region of the detector, which may be located at some specific scattering angle. On the other hand, we are used to thinking of scattered wave functions as having the form $f(\theta; k)\, e^{ikr}/r$, not focused in one particular angle. The only latitude we have is the wave packet leaving the emitter. Is it possible to find an initial wave packet that gathers itself into only some small angle in its final state? This question is most acute for Coulomb scattering which is the nearest thing (neglecting Q.E.D.) to a pure elementary potential and it is least easy to fall back on the argument that most interatomic (say) potentials are only effective interactions built from large

numbers of internal degrees of freedom which may do the tailoring. Rather surprisingly, preliminary calculations with A. Young of The University of Texas suggest that indeed such tailoring is possible and that the remarkable coincidence of the Rutherford scattering cross section for classical and quantum mechanics is what justifies the probability interpretation. But this will be the subject of a later publication.

5. DISCUSSION AND SUMMARY

The problem of quantum measurement theory is that when quantum mechanics is applied to macroscopic systems it seems to predict that there occur superpositions of macroscopically distinct states. Not only are such states absent in our ordinary perception of the world but even if one found a way to justify the (apparent) fact that in an experiment only one of these macroscopic states is selected there seems no mechanism or dynamical law for this selection.

In this paper we have proposed that the measurement problem does not lie in the nature of quantum evolution but in certain assumptions about *a priori* probabilities implicit in the usual transition from the microscopic to the macroscopic. Our approach revolves about a claim, a hypothesis and a postulate.

The *claim* is that there exist special microscopic states of measuring equipment for which the outcome of an experiment is definite, that is, is not a superposition of macroscopically distinct states. This is called a claim because it can be verified—or falsified—within the context of ordinary quantum calculations, at least in principle.

The *hypothesis* is that in all actual experiments, that is, those performed as opposed to contemplated, the initial conditions are among the (possibly rare) special states whose existence is claimed.

The *postulate* is that in the absence of microscopic information about apparatus initial conditions the relative probability of various outcomes is proportional to the relative dimension of the apparatus subspace of special states associated with each outcome.

The object of the paper has been to present and render plausible these theses, not to provide for example a definitive proof that in all measuring apparatus there are special states. In this framework we have also taken liberties in the presentation of incomplete arguments, pointing out however their defects.

The claimed existence of special states is a prediction of sorts. Apart from this claim, what have generally been considered the characteristic properties of a measuring apparatus were amplification and irreversibility. Now we have imposed a further, subtle requirement on the internal state

structure. For our model detection apparatus (Section 2) this requirement was satisfied through the presence of a high degree of degeneracy and uniform coupling of the associated levels to other degrees of freedom. This does not seem unreasonable in an apparatus with highly repetitive arrays, for example the many molecules in a photographic emulsion. Although other sorts of special states may exist, degeneracy could turn out to be a common theme in numerous measurement techniques. This is suggested by the relation between quantum degeneracy and classical regular behavior and the connection of the latter to the appearance of KAM tori in classical dynamical systems.[18–20] Further discussion on special states in scattering theory (hence for decaying particles) is given in Section 4. Recent studies of quantum tunneling[21–23] may provide a more detailed picture of the measurement process—treated entirely quantum mechanically—than has heretofore been known. The SQUIDs and other devices modeled by these theories may provide a clear picture of special states as well as suggest experiments in case there are differences from standard quantum predictions, for example due to problems in producing fully overresponsive states.

By the standards of contemporary hyperbole our central hypothesis could be called a quantum superconspiracy.[9] We offer no explanation for this limitation on our apparent ability to set up arbitrary initial conditions and the associated modification of the usual ideas on the arrow of time. For as we pointed out in Section 3 our theory not only has the usual restrictions on microscopic final states but limits initial states as well.

On this point, however, some historical perspective may be in order. Boltzmann had no idea why final states should be "distinguished," where we use this term to refer to the usual restriction as discussed in Section 3. Speculation in his time suggested that the observed universe was a large fluctuation and the thermodynamic arrow due to a return to equilibrium. And just as I cannot explain why the quantum conspiracy should not sud-

[9] A point that has been raised is that with a superconspiracy *anything* can happen, for example waterfalls that reverse direction each Friday between 2 and 3 p.m. This is in fact true, as it is within the usual statistical mechanics. Consider however why we do not (usually) see reversing waterfalls: For the macroscopically observed waterfall at 1:59, few of its microstates will lead to its reversal at 2:00 (This is not so much "why" as *how* we calculate the likelihood). Waterfalls do not reverse in the present view by exactly the same rationale: For those microstates that are consistent with the 1:59 macrostate *and are special* few lead to reversal at 2 p.m. This is the purport of our earlier discussion on "recovering the ordinary" and is why the presence of the superconspiracy is not generally noticed. This line of reasoning however does lead to an interesting possibility. The quantitative likelihood of certain fluctuations may be different because the available state space is reduced. Hence for borderline-sized fluctuations (not Niagara Falls) the state space restriction may have quantitative implications.

denly terminate and our world fade into a superposition, so they could not justify the belief that memory corresponded to real occurrences, since the entropic cost of an earthquake is surely less than that of its recording in a human brain.

Since Boltzmann's time there has come new information on the universe. Presumably the occurrence of distinguished final states, that is the thermodynamic arrow of time, is a consequence of the expansion of the universe. During expansion regions of phase space open up that describe states that are macroscopically different from earlier configurations. Equilibration does not catch up with expansion, entropy increases and acceptable final states are only those that proceed from earlier states of a more compact universe. How the global process influences local phenomena is a separate question, presumably answerable although attempted answers are often clouded since the questions are not always sharply defined.

In any case, if our central hypothesis is correct its justification may lie in an unanticipated physical phenomenon or perhaps in the nature of perception. We earlier glossed over the problem of defining the term "macroscopic." The subjective or anthropomorphic nature of this term suggests that our perception may filter out non-"definite" outcomes of experiments. At this level of speculation one might tie in to Wigner's view of the role of the observer. (See his first article in Ref. 1) Our point in mentioning these ideas and the quandary of Boltzmann's day is not to point to some particular rationale for our central hypothesis but only to note that failure to see a complete picture does not mean that those features that are visible are false.

Our central hypothesis seems possible only in particular cosmologies, namely, those in which the universe is causally interconnected. This follows from our requirement on the coherence of the global wave function. On this scale the nonlocal effects in the Einstein–Podolsky–Rosen phenomenon seem mild. As mentioned in the Introduction delayed choice experiments, embellishments of the EPR phenomenon, present no problem to us since the existence of choice is an illusion.

Our postulate relating the internal state space of an apparatus to quantum probabilities is another thesis subject to falsification or perhaps to surprising verification. As for our claim on special states, there is at least in principle the possibility of calculating the properties of various sorts of measuring devices. However, it should be noted that this postulate could be false and our claim and central hypothesis true. Then there would be the burden of justifying the recovery of the usual quantum probabilities in some other way.

The traditional test of a theory is an experiment. Unfortunately we do

not know of any observation that could choose between our understanding of quantum mechanics and others. As noted in the Introduction this is because our hypothesis involves no departure from quantum mechanics and it would take a prediction of the other theories that went beyond quantum mechanics before a confrontation could develop. Our postulate on probabilities—since it does go beyond quantum mechanics—could be subject to experimental test. For this one would need precise knowledge of some apparatus with deviations from the usual quantum probabilities occurring perhaps for extremely small versions of the apparatus.

Most of the puzzles of quantum measurement theory involve distance scales of the order 10^{-4} to 10^{-6} cm, the border between macroscopic and microscopic. In contrast to the view implicit in some other theories we do not compromise quantum mechanics on this scale or any other (but with open mindedness on what 15 more orders of magnitude could bring). There is no observational reason, other than the measurement problem, to doubt quantum theory at this level. Now the measurement problem, dealing as it does with the macroscopic world, has both statistical and dynamical assumptions. It is the statistical assumptions we would modify. The usual rule for *a priori* probabilities, that is, give equal weight to all microstates consistent with a macrostate, has never been tested experimentally at the level we have been discussing, although its violation contradicts deeply held intuitions. It is that usual rule that we radically modify through the rejection of all initial states except those yielding definite macroscopic outcomes.

ACKNOWLEDGMENTS

In a previous publication on this subject[8] I thanked Leonard Schulman, Charles Newman, and many other colleagues for help in the formulation and expression of these ideas. Since then I have had the pleasure of further input and argumentation from J. Avron, M. Berry, S. Chakravarty, C. DeWitt, D. Dubin, B. Gaveau, S. Goldstein, M. Gutzwiller, T. Jacobson, S. Kivelson, G. Lasher, J. Lebowitz, B. McCoy, K. Olaussen, C. Radin, P. Seiden, D. Weingarten, A. Young, and the *J. Stat. Phys.* referees. In addition, I have enjoyed the hospitality of several institutions and associated research resources: IBM T. J. Watson Research Center, Drexel University (Louis and Bessie Stein Foundation), The State University of New York at Stony Brook (NSF-DMR-8318051 and NSF-PHY-8109110), The University of Texas at Austin (NSF-PHY-84-04931 and NSF-PHY-82-05717), The University of Paris VI, and the Institute for Theoretical Physics at the University of Trondheim.

APPENDIX A

On the space $L_a \oplus L_b$ the Hamiltonian $H = H_0 + H_1$ is equivalent to the matrix K',

$$K' = \left(\begin{array}{ccc|ccc} E & & 0 & g' & \cdots & g' \\ & \ddots & & \vdots & & \vdots \\ 0 & & E & g' & \cdots & g' \\ \hline g' & \cdots & g' & \omega_1 & & 0 \\ \vdots & & \vdots & & \ddots & \\ g' & \cdots & g' & 0 & & \omega_N \end{array}\right) \tag{A1}$$

where $g' = g/\sqrt{N}$, the upper left block is the $M \times M$ identity times E, the upper right block is an $M \times N$ rectangular matrix every one of whose entries is g', etc.

The first claim made in the text is that $\tilde{\theta}_T$ of Eq. (2.3) is an eigenstate of H. For the matrix representation this amounts to the claim that a column vector whose first M entries add to zero and whose remaining entries are zero is an eigenvector of K'. This is obviously true.

Showing that the remaining dimension in L_a can map almost entirely into L_b is more complicated. Define $e^{(0)}$ to be the column vector whose first M entries are $1/\sqrt{M}$ and remaining entries zero. Let $e^{(k)}$, $k = 1,..., N$ be a vector with a one in the $(M+k)$th row, i.e., the kth row in L_b, zero elsewhere. By direct calculation

$$\begin{aligned} K'e^{(0)} &= Ee^{(0)} + \gamma' \sum_{j=1}^{N} e^{(j)}, \qquad \gamma' \equiv g'\sqrt{M} = \frac{1}{\sqrt{N}}\, g\sqrt{M} \equiv \frac{1}{\sqrt{N}}\,\gamma \\ K'e^{(j)} &= \gamma' e^{(0)} + \omega_j e^{(j)}, \qquad j = 1,..., N \end{aligned} \tag{A2}$$

With no loss of generality we take $E = 0$ since only an irrelevant phase may be changed. From (A2) it follows that the modes of L_a corresponding to $\tilde{\theta}_T$ are completely decoupled and the matrix K' can be replaced by the $(N+1) \times (N+1)$ matrix K

$$K = \begin{pmatrix} 0 & \gamma' & \gamma' & \cdots & \gamma' \\ \gamma' & \omega_1 & & & \\ \gamma' & & \ddots & \bigcirc & \\ \vdots & \bigcirc & & & \\ \gamma' & & & & \omega_N \end{pmatrix} \tag{A3}$$

We want the time dependence of a state with initial condition $e^{(0)}$. In particular, we want to justify our claim, made in the text, that at certain times given approximately by $T=(n+\frac{1}{2})\,\pi/g\sqrt{M}\quad n=0,1,\ldots$ the system lies almost entirely in L_b. To do this we shall show that

$$\delta(T)\equiv e^{(0)\dagger}\exp(-iKT)\,e^{(0)}=O(\omega_0^2/g^2M) \tag{A4}$$

where $\omega_j\in[-\omega_0,\omega_0]$. The "narrow band" feature of the phonons is taken to mean $\omega_0^2\gtrsim g^2$ and therefore $\omega_0^2\ll Mg^2$. We shall also assume that N is even and that the levels ω_j are symmetric about zero. We shall establish (A4) using an eigenvector expansion for $\exp(-iKT)$.

Let v be an eigenvector of K with components $x_0, x_1,\ldots, x_N$ and eigenvalue λ. Then $Kv=\lambda v$ implies

$$\begin{aligned}\gamma'\sum_{j=1}^{N}x_j&=\lambda x_0\\ x_0\gamma'+\omega_k x_k&=\lambda x_k,\qquad k=1,\ldots,N\end{aligned} \tag{A5}$$

Combining, we obtain a consistency condition determining λ

$$\lambda=\gamma^2\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\lambda-\omega_j} \tag{A6}$$

and the form of the eigenvector

$$\begin{aligned}x_k&=\frac{\gamma}{\sqrt{N}}\frac{x_0}{\lambda-\omega_k},\qquad k=1,\ldots,N\\ x_0&=\left[1+\gamma^2\frac{1}{N}\sum_{j=1}^{N}\left(\frac{1}{\lambda-\omega_j}\right)^2\right]^{-1/2}\end{aligned} \tag{A7}$$

the last expression fixed by the normalization $v^\dagger v=1$. Defining $f(\lambda)=N^{-1}\sum_{j=1}^{N}(\lambda-\omega_j)^{-1}$, we note the following features of the curve $f(\lambda)$ versus λ. For large positive (negative) λ it is positive (negative) and drops off like const/λ. For λ immediately to the right (left) of any ω_j it is large and positive (negative). The solutions of (A6) are the intersections of $f(\lambda)$ with the line λ/γ^2. There will be $N-1$ such intersections in the interval $[-\omega_0,\omega_0]$, one between each pair of ω_j's, and a pair of symmetrically disposed roots far from this interval for $\gamma\gg\omega_0$. Symmetry of the set of ω_j's under reflection about zero implies the same property for the set of λ's. We shall not need further information on the eigenvalues clustered in $[-\omega_0,\omega_0]$ but we shall determine properties of the large $|\lambda|$ solution.

To lowest order in ω_0/γ the ω's in the denominator of f can be neglected [this can be seen by defining a new unknown as λ/γ and deriving its

equation from (A6)]. This gives $\lambda = \pm\gamma$. Correction to this expression depends on the details of the density of states. Since $-\lambda$ is also a solution of (A6) it follows that $\lambda = -\gamma^2 f(-\lambda)$ is also true. Adding this to (A6) and dividing by 2λ yields $1 = \gamma^2 \sum(\lambda^2 - \omega_j^2)^{-1}$ or

$$\frac{\lambda^2}{\gamma^2} = 1 + \frac{1}{N}\sum \frac{\omega_j^2}{\lambda^2} + \frac{1}{N}\sum \left(\frac{\omega_j}{\lambda}\right)^4 \frac{1}{\lambda - \omega_j} \tag{A8}$$

Neglecting the second sum, we set $N^{-1}\sum \omega_j^2 = n_0\omega_0^2$ where the constant n_0 is of order unity and depends on the density of states. The resulting quadratic equation yields

$$\begin{aligned} \lambda^2 &= \gamma^2 + \frac{\gamma^2}{2}\left[\left(1 + \frac{4n_0\omega_0^2}{\gamma^2}\right)^{1/2} - 1\right] \cong \gamma^2 + n_0\omega_0^2 \\ \lambda &\cong \gamma + \frac{1}{2}\frac{n_0\omega_0^2}{\gamma} = \gamma\left[1 + O\left(\frac{\omega_0^2}{\gamma^2}\right)\right] \end{aligned} \tag{A9}$$

where only one of the roots of the original quadratic is taken since the other yields a value of λ of small absolute value, contradicting the assumptions used in its derivation. (The foregoing manipulations can be carried out in the continuum approximation to all orders in ω_0/γ for a constant density of states. Solving the resulting transcendental equation gives (A9) as the first-order correction to $\lambda = \gamma$.)

The $N+1$ eigenvalues and eigenvectors are labeled $\lambda_n, v^{(n)}$ and we define the matrix V to have $v^{(n)}$ as its nth column. Thus

$$\begin{aligned} &V_{mn} = v_m^{(n)}, \qquad v_m^{(n)} = m\text{th component of } v^{(n)} \\ &V^T = V^{-1} \\ &WV = VW_D \qquad \text{where } W_D = \operatorname{diag}(\lambda_0, \lambda_1, \ldots, \lambda_N) \end{aligned} \tag{A10}$$

The second equality follows from the orthonormality of the eigenvectors of the real, Hermitian matrix K whose eigenvalues are nondegenerate. To evolve $e^{(0)}$ in time we express it in terms of $v^{(n)}$. Using $V^{-1} = V^T$ this expression is

$$e^{(0)} = \sum_{n=0}^{N} v_0^{(n)} v^{(n)} \tag{A11}$$

where $v_0^{(n)}$ is the same as x_0 in (A7) with λ_n replacing λ. Under time evolution $v^{(n)} \to \exp(-i\lambda_n T)\, v^{(n)}$ and the expression $\delta(T)$ of Eq. (A4) becomes

$$\delta(T) = \sum_{n=0}^{N} (v_0^{(n)})^2\, e^{-i\lambda_n T} \tag{A12}$$

Let the two large eigenvalues of K be $\bar{\lambda}$ (>0) and $-\bar{\lambda}$ (the numbering is irrelevant). The respective zero components of their eigenvectors are equal (by symmetry of the ω's) so that for times T such that $\cos \bar{\lambda} T = 0$ their combined contribution to the sum in (A12) vanishes. This will be taken to be the time needed to obtain "total" absorption, for which the expression given in Eq. (2.5) is an approximation. Let $\sum'$ represent the sum over eigenvalues other than $\bar{\lambda}$ and $-\bar{\lambda}$ and let δ designate $|\delta(T)|$ at one of the times for which $\cos \bar{\lambda} T = 0$. Then

$$\delta = \left| \sum{}' (v_0^{(n)})^2 \, e^{-i\lambda_n T} \right| \leqslant \sum{}' (v_0^{(n)})^2 \tag{A13}$$

By $(VV^T)_{00} = 1$ the sum in (A13) is one minus the sum over $\bar{\lambda}$ and $-\bar{\lambda}$ alone. Since the two latter terms are equal we have

$$\delta \leqslant 1 - 2\left[1 + \gamma^2 \frac{1}{N} \sum_{k=1}^{N} \frac{1}{(\bar{\lambda} - \omega_k)^2} \right]^{-1} \tag{A14}$$

where the explicit form for the zeroth component is used. Let

$$\delta' = \gamma^2 \frac{1}{N} \sum_{k=1}^{N} \frac{1}{(\bar{\lambda} - \omega_k)^2} - 1 \tag{A15}$$

Then if we show $\delta' = O(\omega_0^2/\gamma^2)$ it will follow that δ is of the same (small) order.

Again by the symmetry of the ω_j's the sum in (A14) is the same for $\bar{\lambda}$ and $-\bar{\lambda}$ so that

$$\begin{aligned} \delta' &= \frac{1}{2} \gamma^2 \frac{1}{N} \sum \left[\left(\frac{1}{\bar{\lambda} - \omega_j} \right)^2 + \left(\frac{1}{\bar{\lambda} + \omega_j} \right)^2 \right] - 1 \\ &= \frac{\gamma^2}{\bar{\lambda}^2} \frac{1}{N} \sum (1 + \omega_j^2/\bar{\lambda}^2)(1 - \omega_j^2/\bar{\lambda}^2)^{-2} - 1 \end{aligned} \tag{A16}$$

Since $1/N$ times the sum gives $1 + O(\omega_0^2/\bar{\lambda}^2)$ and since $\bar{\lambda}^2/\gamma^2$ is $1 + O(\omega_0^2/\gamma^2)$ it follows that $\delta' = O(\omega_0^2/\gamma^2)$ or $\delta' = O(\omega_0^2/g^2 M)$, with the same consequence for δ.

Finally we wish to show that for undistinguished initial conditions the probability for detection in this cluster is $1/M$. As an initial state in L_a we take

$$\theta = \phi^\dagger \sum \beta_k a_k^\dagger |0\rangle \qquad \sum |\beta_k|^2 = 1 \tag{A17}$$

with either a uniform distribution in the β's or a particular β_k unity, the others zero. Reexpress θ in terms of $e^{(0)}$ and the other (totally unrespon-

sive) vectors in L_a. The latter stay in L_a while $e^{(0)}$ goes almost completely to L_b. The probability of absorption will therefore be the absolute value squared of the coefficient of $e^{(0)}$ in an expansion of θ. Therefore

$$p_A = \text{probability of absorption} = \frac{1}{M}\left|\sum \beta_k\right|^2 \tag{A18}$$

If $\beta_k = \delta_{kj}$, $j = 1,..., M$ then p_A is clearly $1/M$. On the other hand, if the β_k are uniformly distributed and uncorrelated then from the normalization in (A17) we again have the expected value of p_A being $1/M$.

REFERENCES

1. J. A. Wheeler and W. H. Zurek, eds., *Quantum Theory and Measurement* (Princeton University Press, Princeton, New Jersey, 1983).
2. K. Gottfried, *Quantum Mechanics* (Benjamin, New York, 1966); especially Section 20.3.
3. L. E. Ballentine, *Rev. Mod. Phys.* **42**:358 (1970).
4. P. Edwards and A. Pap, *A Modern Introduction to Philosophy* (Macmillan, New York, 1973).
5. D. J. O'Connor, *Free Will* (Anchor Doubleday, Garden City, New York, 1971).
6. D. C. Dennett, *Elbow Room* (MIT Press, Cambridge, Massachusetts, 1984).
7. A. Komar, *Phys. Rev.* **126**:365 (1962).
8. L. S. Schulman, *Phys. Lett.* **102A**:396 (1984).
9. H. Everett III, *Rev. Mod. Phys.* **29**:454 (1957); reprinted in Ref. 1.
10. P. K. Kabir, *The CP Puzzle: Strange Decays of the Neutral Kaon* (Academic Press, New York, 1968) Appendix A.
11. L. S. Schulman, R. G. Newton, and R. Shtokhamer, *Philos. Sci.* **42**:503 (1975).
12. M. Kac, *Probability and Related Topics in Physical Sciences* (Interscience, New York, 1959) Sections 14 and 15.
13. L. S. Schulman, *J. Stat. Phys.* **16**:217 (1977).
14. V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics* (Benjamin, New York, 1968).
15. L. S. Schulman and R. Shtokhamer, *Int. J. Theor. Phys.* **16**:287 (1977); see also the Appendix to Ref. 13.
16. L. S. Schulman, *Techniques and Applications of Path Integration* (Wiley, New York, 1981).
17. H. Goldstein, *Classical Mechanics*, 2nd ed. (Addison-Wesley, Reading, Massachusetts, 1980).
18. M. V. Berry, Semiclassical mechanics of regular and irregular motion, in *Chaotic Behavior of Deterministic Systems*, G. Iooss, R. Helleman and R. Stora, editors (North-Holland, Amsterdam, 1983).
19. M. V. Berry and M. Tabor, *Proc. Soc.* **A356**:375 (1977).
20. A. S. Wightman, in *Perspectives in Statistical Physics*, H. J. Reveche, ed. (North-Holland, Amsterdam, 1981).
21. A. O. Caldeira and A. J. Leggett, *Phys. Rev. Lett.* **46**:211 (1981).
22. S. Chakravarty, Macroscopic Quantum Coherence in Superconducting Interference Devices, presented at conference on "Fundamental Questions in Quantum Mechanics," SUNY at Albany, 1984.
23. A. J. Leggett, in Proc. Int. Symp. Foundations of Quantum Mechanics, Tokyo, 1983.