International Journal of Theoretical Physics, Vol. 18, No. 7, 1979

Toward Explaining Why Events Occur

Philip Pearle

Hamilton College, Clinton, New York 13323

Received April 26, 1979

The possibility is raised of adding a randomly fluctuating interaction term to the Schrödinger equation, so that the new equation "reduces" the state vector. The exact event that occurs is predicted by the equation and depends upon the precise time dependence of the interaction term. The uncertainty in nature is attributed to the random behavior of this term. A class of such terms is found. This class includes terms whose nonlinear dependence on the wave function is identical to that of terms introduced in a previous paper for a similar purpose. In the previous paper, the exact event predicted depends upon the initial phase factors in the superposition making up the state vector: the uncertainty in nature is attributed to random initial phase factors. Another derivation of the results in the previous paper is given in an appendix: the calculations in that paper and in this appendix are of second order in perturbation theory. On the other hand, the calculations in the present paper are exact. A possible answer is given to the question, raised in the previous paper, of the nature of the "observable" states to which the state vector reduces.

1. INTRODUCTION

Although quantum theory is the most successful predictive theory in science, it nevertheless is incapable of explaining—let alone predicting—the individual events that occur in every quantum experiment (the time of decay of a single particle, the angle into which a particular particle projectile scatters, etc). The state vector (wave function) of quantum theory must be interpreted as corresponding to an ensemble of identical physical systems (Jauch, 1971; Belinfante, 1975).

Suppose one seeks an alternative theory that will explain the occurrence of an individual event. Rather than modifying quantum theory by introducing a theory with additional variables (Belinfante, 1973), it seems simplest to follow Schrödinger's original vision of the wave function as the description of an individual system. But then the dynamics of the wave

0020-7748/79/0700-0489$03.00/0

function must be modified: that is, Schrödinger's own equation must be altered to achieve Schrödinger's interpretation of the wave function.

To see this, consider the state vector $|\psi,t\rangle$ corresponding to the physical system consisting of an apparatus making a measurement on a single microscopic system. At a time when the measurement has been completed, the state vector can be written

$$|\psi,t\rangle = \sum_n a_n(t)|\phi_n(t)\rangle \tag{1.1}$$

where the $|\phi_n(t)\rangle$ correspond to different macroscopic outcomes of the experiment. According to Schrödinger's equation, in general, there are many a_n's for which $0<|a_n(t)|^2<1$. However, if the state vector is to describe a single system, we must have at a sufficiently large time, for some m,

$$|a_m(t)|^2 \to 1, \qquad |a_n(t)|^2 \to 0, \qquad n \neq m \tag{1.2}$$

As Schrödinger (1935) himself pointed out most graphically with his "cat paradox," a single macroscopic system never ends up in a physical state that can be described by a superposition of macroscopically different states; it must end up in a physical state that can be described by a state vector such as $|\psi,t\rangle = |\phi_m\rangle$.

Therefore, in order to be able to describe a single physical system by a state vector, we seek a modification of Schrödinger's equation that works as follows.

We assume that a single microscopic system *together* with its environment up to, and possibly including, the rest of the universe, is the physical system appropriately described by a state vector [this hypothesis has also been invoked in the case of quantum theory (Everett, 1957; Pearle, 1967)]. The modified Schrödinger equation is to be designed so that, like Schrödinger's equation, it possesses the following property.

Property 0. The coefficients in the superposition (1.1) always satisfy $\Sigma_n|a_n(t)|^2 = 1$.

The equation is to act like the ordinary Schrödinger equation until, say time $t=0$, when the state vector has the form (1.1), which one may think of as a "linear superposition of pointer positions." Immediately following this juncture, the "pointer positions fight it out." That is to say, the modified Schrödinger equation possesses the following property.

Property 1. The coefficients in the superposition (1.1) end up as in (1.2), with all vanishing except one.

These properties are still insufficient. In order to make as good predictions of the outcomes of experiments as does quantum theory, the modified Schrödinger equation must possess a further property. It must be possible to obtain many different solutions of this equation (each possessing property 1), such that the set of solutions possesses the following property.

Property 2. The fraction of the total number of solutions for which the coefficient $|a_m(t)|^2 \to 1$, must be equal to $|a_m(0)|^2$.

(We are supposing that the usual quantum mechanical description is essentially complete at time $t=0$, so that $|a_m(0)|^2$ is the quantum theory prediction of the probability of the mth outcome of the experiment.)

We shall call an equation possessing properties 0, 1, and 2 a "reducing Schrödinger equation," or just a "reducing equation."

In a previous paper (Pearle, 1976) (hereafter referred to as I), the author presented a simple sufficiency test for an equation to be a reducing equation. Let $x_n(t) \equiv |a_n(t)|^2$, and denote by $\langle\ \rangle$ the average over the set of solutions of the equation. If, in addition to property 0, the following conditions hold:

$$\langle x_n(t) \rangle = x_n(0), \qquad 0 \leqslant t < \infty \tag{1.3a}$$

$$\langle x_n(t) x_m(t) \rangle \underset{t \to \infty}{\to} 0, \qquad \text{all } n \neq m \tag{1.3b}$$

then the equation is a reducing equation.[1]

Here is why equations (1.3) imply properties 1 and 2. Because each $x_n(t)$ is nonnegative, equation (1.3b) requires that for each solution in the set of solutions, at least one member of each pair x_n, x_m asymptotically approach zero. But this means that for each solution, all x_n except possibly one must asymptotically approach zero. Because of property 0, indeed one x_n must asymptotically approach 1. Thus property 1 is guaranteed by equation (1.3b). [It is interesting that, although equation (1.3b) is a statement about average behavior of solutions, it is capable of supplying information about an individual solution.]

Now we may apply equation (1.3a) as $t \to \infty$, obtaining

$$1 \times \text{Prob}\left[x_n(t) \underset{t \to \infty}{\to} 1 \right] + 0 \times \text{Prob}\left[x_n(t) \underset{t \to \infty}{\to} 0 \right] = x_n(0) \tag{1.4}$$

[1]The choice of $t \to \infty$ defining asymptotic behavior is made for mathematical convenience only. The reduction process should essentially be completed, for almost all solutions, in a finite amount of time. In fact, one hopes that a characteristic reduction time can be experimentally determined.

from which we find that the fraction (probability) of solutions for which $x_n(t)$ asymptotically approaches 1 is $x_n(0)$. This is property 2. [We remark that although all we have used in this argument is the asymptotic behavior of (1.3a), we need the stronger form of this equation because it is difficult to construct a reducing equation that provides the appropriate asymptotic behavior unless it satisfies equation (1.3a) at all times.]

How does one go about constructing a reducing Schrödinger equation? It seems simplest to add an extra term to the ordinary Schrödinger equation. This term may be thought of as a new kind of interaction, a nonlocal interaction between "pointer positions." This interaction has two important aspects. First, the interaction "matrix elements" are supposed to be negligibly small between states that are macroscopically indistinguishable, but large enough to dominate the dynamics when the states are macroscopically distinguishable. Second, the term's functional dependence on $a_n(t)$ must be responsible for the reducing behavior. It is this second aspect that we addressed in I, and continue to address in this paper. (The first aspect, which is certainly of great importance to a complete theory, will be briefly commented upon in our concluding remarks.)

We have said that a reducing Schrödinger equation must have many different solutions. Each solution predicts a definite outcome for the experiment that is being described. The inability of physicists to predict the precise outcome of an experiment is attributed to our inability to experimentally fix the variables responsible for determining the different solutions. What are these variables?

In I, these variables were taken to be initial conditions, specifically the initial phase factors θ_n [$a_n \equiv (x_n)^{1/2} \exp i\theta_n$] were responsible for determining the final state vector.

In this paper, we explore another mechanism. The idea is that the solution of the reducing Schrödinger equation is different each time we solve it because the equation is different each time we solve it. That is, the interaction term is taken to be proportional to a random function of time. If we knew the precise time dependence of this function we could predict with certainty the result of the experiment.

2. SUMMARY OF RESULTS

In order to more precisely describe the content of this paper, it is useful to review the results obtained in I. Consider the ordinary Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi, t\rangle = H|\psi, t\rangle + A|\psi, t\rangle \tag{2.1}$$

$|\psi,t\rangle$ has an expansion (1.1) in basis states $|\phi_n(t)\rangle$ which are assumed to be known and to satisfy

$$i\hbar\frac{d}{dt}|\phi_n(t)\rangle=\{H_0-\hbar\omega_n\}|\phi_n(t)\rangle, \qquad \omega_n\equiv\langle\phi_n|H_0|\phi_n\rangle \tag{2.2}$$

This equation, which implies that $\langle\phi_n(t)|d|\phi_n(t)\rangle/dt=0$, settles the question of how much phase factor in the product $a_n|\phi_n\rangle$ in equation (1.1) belongs to a_n, and how much belongs to $|\phi_n\rangle$. Equation (2.2) makes $|\phi_n(t)\rangle$ time independent if it is an eigenstate of H_0.

Define the projection operator $P_n(t)\equiv|\phi_n(t)\rangle\langle\phi_n(t)|$, and premultiply equation (2.1) by $\langle\psi,t|P_n(t)$:

$$i\hbar\left\langle\psi\left|P_n\frac{d}{dt}\right|\psi\right\rangle=\langle\psi|P_nH|\psi\rangle+\langle\psi|P_nA|\psi\rangle \tag{2.3}$$

No information is lost by doing this: if one divides equation (2.3) by $\langle\psi,t|\phi_n(t)\rangle$, one obtains Schrödinger's equation expressed in the $\langle\phi_n(t)|$ basis. Now take equation (2.3) and rewrite it, but replace the last term by its complex conjugate:

$$i\hbar\left\langle\psi\left|P_n\frac{d}{dt}\right|\psi\right\rangle=\langle\psi|P_nH|\psi\rangle+\langle\psi|AP_n|\psi\rangle \tag{2.4}$$

Equation (2.4) is no longer an equation linear in $|\psi\rangle$, for if we divide (2.4) by $\langle\psi|\phi_n\rangle$, this factor does not disappear, as we obtain

$$i\hbar\left\langle\phi_n\left|\frac{d}{dt}\right|\psi\right\rangle=\langle\phi_n|H|\psi\rangle+\sum_m\langle\psi|\phi_m\rangle\langle\phi_m|A|\phi_n\rangle\frac{\langle\phi_n|\psi\rangle}{\langle\psi|\phi_n\rangle} \tag{2.5}$$

Nonetheless, equation (2.4) possesses property 0, for if (2.3), (2.4) are summed over n, the resulting expressions, which imply $d\langle\psi|\psi\rangle/dt=0$, are identical.

It was shown in I that equation (2.4) [or equation (2.5)] is a reducing Schrödinger equation under certain assumptions. [Actually, a wider class of reducing equations was considered in I, but for reasons given there, and for some to be given here, (2.4) seems to be the simplest choice.] It was assumed that the effect of the nonlinear interaction term could be adequately treated in second-order perturbation theory. Then, by two different approaches (one due to Prigogine, the other due to Markov-Chandrasekhar), it was shown that one could obtain a partial differential (Fokker–Planck) equation to describe the probability density function $\rho(x_1,x_2,\ldots,t)$

for the variables x_n, in the solution set of equation (2.4) with random initial phase factors:

$$\frac{\partial \rho}{\partial t} = \sum_{n,m} |\alpha_{nm}|^2 \left(\frac{\partial}{\partial x_n} - \frac{\partial}{\partial x_m} \right)^2 x_n x_m \rho, \qquad t \geqslant 0 \tag{2.6}$$

(here $\alpha_{nm} \sim \langle \phi_n | A | \phi_m \rangle$). In Appendix A we present yet another derivation of this result.

It is easily seen that if the initial conditions imply $\Sigma x_n(0) = 1$, then

$$\rho(t) \sim \delta\left(1 - \sum_n x_n\right) \tag{2.7a}$$

is satisfied by the solution of equation (2.6). This entails property 0.

Furthermore, if we multiply equation (2.6) by x_n or by $x_n x_m$, and integrate over all x_i from 0 to 1, we obtain, respectively,

$$\frac{d}{dt} \langle x_n \rangle = 0 \tag{2.7b}$$

$$\frac{d}{dt} \langle x_n x_m \rangle = -4 |\alpha_{nm}|^2 \langle x_n x_m \rangle, \qquad n \neq m \tag{2.7c}$$

(where $\langle x_n \rangle \equiv \int d\mathbf{x} \rho x_n$, $\langle x_n x_m \rangle \equiv \int d\mathbf{x} \rho x_n x_m$). Clearly, equations (2.7b) and (2.7c) imply the conditions (1.3). Thus, properties 1 and 2 are satisfied.

It should be noted that this proof (see Appendix A) requires many states $|\phi_n\rangle$ closely spaced in energy, since otherwise, as in quantum second-order perturbation theory, the amplitudes a_n merely oscillate. This means, for example, that the solutions of equation (2.4) for a two-dimensional Hilbert space do not exhibit reduction behavior. It also should be mentioned that the nature of the "observable" states $|\phi_n(t)\rangle$, to one of which the state vector reduces, is a question separate from the question of the mechanism of the reduction process, and will be addressed in Section 5.

We now summarize the results of this paper. We have investigated whether a reducing Schrödinger equation can be found in the form

$$i\hbar \left\langle \phi_n \middle| \frac{d}{dt} \middle| \psi \right\rangle = \langle \phi_n | H | \psi \rangle + \sum_m A_{mn} \dot{B}_{mn} F(\langle \phi_m | \psi \rangle, \langle \phi_n | \psi \rangle) \tag{2.8}$$

In equation (2.8), $\dot{B}_{mn}$ is an element of a Hermitian matrix composed of random functions of time. B_{mn} is a complex Brownian motion, and $\dot{B}_{mn}$ is a

complex so-called "white noise" (these concepts will be introduced more carefully in the next section). A_{mn} is an element of a Hermitian matrix, the nonlocal interaction matrix mentioned earlier, that is supposed to be large only when $|\phi_m\rangle, |\phi_n\rangle$ are macroscopically distinguishable.

Our task is to find what form the (so far) arbitrary function F must take in order for (2.8) to be a reducing equation. As it stands, strictly speaking, equation (2.8) does not exist, since the derivative of Brownian motion—white noise—does not exist. However, there is a well-known procedure for giving a meaning to equation (2.8), where $\dot{B}_{mn}$ is regarded as the limit of a sequence of well-defined functions. This leads us to convert equation (2.8) to a set of (so-called) stochastic differential equations. Next we follow a well-known procedure for finding the Fokker–Planck equation describing the probability density function for the solutions of the stochastic equations. We may then use the Fokker–Planck equation to impose the conditions (1.3) which ensure that equation (2.8) will be a reducing equation. This leads to a partial differential equation that F must satisfy. We find a class of solutions for F, including

$$F = \langle\psi|\phi_m\rangle\langle\phi_n|\psi\rangle / \langle\psi|\phi_n\rangle \tag{2.9}$$

which we believe is the simplest choice. The Fokker–Planck equation, when integrated over phase angles, becomes equation (2.6).

In short, we find that equation (2.4) is also a reducing equation when the initial phases are not random, but instead the operator A is a random operator, with matrix elements of the form $A_{mn}\dot{B}_{mn}$. Which mechanism is preferable to account for the uncertainty in nature: random initial conditions or a random interaction? Physically speaking, at present it is hard to say (but see Section 5). One may hope that a reducing equation like (2.4) might arise, perhaps as an approximation, in the context of a larger theory (e.g., a quantum theory of gravity), in which case the overriding theory would answer this question. However, mathematically speaking, the theory presented in this paper is to be preferred on two counts. First, the theory presented here is applicable to a two-dimensional (or any-dimensional) Hilbert space. Second, this theory is exact, not a second-order perturbation approximation.

Our discussion proceeds as follows. In the next section the needed statistical concepts and results are presented. In Section 4 they are applied to obtain a reducing Schrödinger equation. In Section 5 we suggest a method of obtaining the observable basis $|\phi_n\rangle$ and the Hamiltonian H_0 governing its time evolution. The last section raises some questions still to be answered before this can be considered a satisfactory theory.

3. STOCHASTIC DIFFERENTIAL EQUATIONS

This section contains an informal introduction to the definitions and theorems needed in the next section. Rigorous proofs, or references to those proofs, can be found in the excellent text by Wong (Wong, 1971; see also Wong and Zakai, 1965; McShane, 1960), or elsewhere (Friedman, 1975; McKean, 1969; Ash and Gardner, 1975).

A *stochastic process* is a function of two variables, a coordinate (Ω) in a probability space, and time (t). It is called a process because it proceeds in time. It is called stochastic to distinguish it from a causal process (a fancy name for a function that depends upon time only).

We shall take the name "*Brownian motion*" to be synonomous with the Wiener stochastic process $B(t)$ (the dependence on the variable Ω is suppressed), which is completely defined by the following conditions:

(i) Each "sample function" $B(t)$ is a real continuous function of t.

(ii) $B(0)=0$.

(iii) $B(t_2)-B(t_1)$ has a Gaussian (normal) probability distribution with mean 0 and variance $\sigma^2|t_2-t_1|$. [Note that this completely describes the probability distribution of $B(t)$.]

(iv) $B(t_N)-B(t_{N-1})$, $B(t_{N-1})-B(t_{N-2}),\ldots,B(t_2)-B(t_1)$ are statistically independent if $t_N \geqslant t_{N-1} \geqslant t_{N-2} \cdots \geqslant t_2 \geqslant t_1$.

Any expectation value of products of B's can be found using (i)–(iv): for example, if $t_2 \geqslant t_1$:

$$\langle B(t_2)B(t_1)\rangle = \langle [B(t_2)-B(t_1)][B(t_1)-B(0)]\rangle + \langle [B(t_1)-B(0)]^2\rangle$$
$$= 0+\sigma^2 t_1 = \sigma^2 \min(t_2,t_1) \tag{3.1}$$

(here $\langle\ \rangle$ represents the average over the probability space).

It is helpful to know that Brownian motion can be represented as the limit of a Fourier series of period T (Wang and Uhlenbeck, 1945)

$$B(t)=\lim_{T\to\infty}\frac{1}{T}\sum_{k=0}^{\infty}\frac{1}{\left(\frac{2\pi k}{T}\right)}\left\{a_k\left[1-\cos\left(\frac{2\pi kt}{T}\right)\right]+b_k\sin\left(\frac{2\pi kt}{T}\right)\right\} \tag{3.2}$$

where each a_k, b_k are independent random variables having a Gaussian distribution, zero mean, and variance $2\sigma^2 T$. One can formally take the limit in (3.2) to obtain

$$B(t)=\frac{1}{2\pi}\int_0^\infty \frac{d\omega}{\omega}\left[a(\omega)(1-\cos\omega t)+b(\omega)\sin\omega t\right] \tag{3.3a}$$

$$\langle a(\omega)\rangle = \langle b(\omega)\rangle = 0, \qquad \langle a(\omega')a(\omega)\rangle = \langle b(\omega')b(\omega)\rangle = 4\pi\sigma^2\delta(\omega'-\omega) \tag{3.3b}$$

Any expectation values of products of B's can alternatively be found using the Fourier representation.

The derivative of equation (3.3a) is formally

$$\dot{B}(t)=\frac{1}{2\pi}\int_0^\infty d\omega\left[a(\omega)\sin\omega t+b(\omega)\cos\omega t\right] \tag{3.4}$$

$\dot{B}(t)$ is called white noise. One might think of it as a Gaussian random process with zero mean and infinite variance, or as a process in whose frequency spectrum all frequencies are equally likely—except that, strictly speaking, $\dot{B}(t)$ does not exist (Brownian motion is nondifferentiable). White noise may best be considered as the limit of a process with a cutoff frequency spectrum [Equation (3.4) with a finite upper limit on the integral] as the cutoff increases. It is used in physical calculations as an idealization (since any physical random quantity has a high-frequency cutoff), largely because of its elegant mathematical qualities. Like the Dirac δ function, it is really only used in integrals. Indeed, it is intimately related to the δ function, for from (3.4), (3.3b) we obtain formally

$$\langle\dot{B}(t_2)\dot{B}(t_1)\rangle=\frac{1}{(2\pi)^2}4\pi\sigma^2\int_0^\infty d\omega\cos\omega(t_2-t_1)=\sigma^2\delta(t_2-t_1) \tag{3.5}$$

which leads to equations like

$$\left\langle\int_a^b f(t_2)dB(t_2)\int_a^b g(t_1)dB(t_1)\right\rangle=\sigma^2\int_a^b f(t)g(t)\,dt \tag{3.6}$$

which can be rigorously justified.

We are interested in solving differential equations involving Brownian motion that are linear in white noise. A simple example is

$$\frac{dX(t)}{dt}=F(B(t))\dot{B}(t) \tag{3.7}$$

One might expect that equation (3.7) can be solved by writing

$$dX(t)=F(B(t))dB(t) \tag{3.8}$$

and performing the integral. Surprisingly, this is not correct. That is, equations (3.7) and (3.8), as we shall interpret them, are not generally equivalent.

Let us start with the differential equation (3.8), which is not well defined: do we mean by (3.8)

$$\Delta X_1(t)=F(B(t))\left[B(t+\Delta t)-B(t)\right] \tag{3.9a}$$

or

$$\Delta X_2(t) = F(B(t+\Delta t))\left[B(t+\Delta t) - B(t)\right] \tag{3.9b}$$

(in the limit as $\Delta t \to 0$), or some other definition where the argument of F is evaluated at some intermediate time between t and $t+\Delta t$? To see that these are inequivalent definitions, set $F = B$, and use (i)–(iv) to evaluate the expectation value of equations (3.9):

$$\Delta\langle X_1(t)\rangle = \langle B(t)\left[B(t+\Delta t) - B(t)\right]\rangle$$

$$= \langle\left[B(t) - B(0)\right]\left[B(t+\Delta t) - B(t)\right]\rangle = 0 \tag{3.10a}$$

$$\Delta\langle X_2(t)\rangle = \langle B(t+\Delta t)\left[B(t+\Delta t) - B(t)\right]\rangle = \langle\left[B(t+\Delta t) - B(t)\right]^2\rangle$$

$$= \sigma^2 \Delta t \tag{3.10b}$$

Of course, the same results can be obtained using (3.3) and (3.4):

$$\Delta\langle X_{\frac{1}{2}}^{1}(t)\rangle \equiv \lim_{\varepsilon\to 0}\langle B(t \mp \varepsilon)\dot{B}(t)\Delta t\rangle$$

$$= \lim_{\varepsilon\to 0}\left(\frac{\sigma^2\Delta t}{\pi}\right)\int_0^\infty \frac{d\omega}{\omega}(\sin\omega t \mp \sin\omega\varepsilon) = \begin{cases} 0 \\ \sigma^2\Delta t \end{cases} \tag{3.11}$$

If $B(t)$ was a Riemann integrable function, the forward difference approximation (3.9a) or the backward difference approximation (3.9b) would lead to the same results in the limit $\Delta t \to 0$. In this case they do not. In our example, although we have not yet figured out what X_1 and X_2 are, by integrating (3.10) or (3.11) we see that their expectation values satisfy

$$\langle X_1(t)\rangle - \langle X_1(0)\rangle = 0 \tag{3.12a}$$

$$\langle X_2(t)\rangle - \langle X_2(0)\rangle = \sigma^2 t \tag{3.12b}$$

Equation (3.12a), called the *Martingale* property, also holds for Equation (3.9a), where F is an *arbitrary* function of Brownian motion [using property (iv)]. Because of the simplicity of the Martingale property, it is useful to define a *stochastic integral* using forward differences:

$$\int_b^a F(t, Z(t))dB(t) \equiv \lim_{\Delta t\to 0}\sum_{n=0}^{(b-a)/\Delta t} F(t_n, Z(t_n))\left[B(t_n+\Delta t) - B(t_n)\right] \tag{3.13}$$

where $t_n \equiv a + n\Delta t$, and $Z(t_n)$, which depends in any way on Brownian motion at times prior to (or equal to) t_n, is called a *nonanticipating function.*

Having defined what is meant by a stochastic integral, we can proceed to integrate (3.8), seeking a function $X(t)$ satisfying

$$X(t) - X(0) = \int_0^t F(B(t))dB(t) \tag{3.14}$$

Equation (3.8) is interpreted as equivalent to equation (3.14). The problem of how to integrate (3.14) was solved by Itô (who, incidentally, raised the problem by defining the stochastic integral). Here is a heuristic derivation of *Itô's differentiation rule.*

Consider a function $Y(t, B(t))$. Using Taylor's expansion we obtain,

$$\begin{aligned} dY(t) &\equiv Y(t + dt, B(t) + dB) - Y(t, B(t)) \\ &= (\partial_t Y)\,dt + (\partial_B Y)\,dB + \tfrac{1}{2}(\partial_t^2 Y)(dt)^2 + (\partial_{tB}^2 Y)\,dt\,dB + \tfrac{1}{2}(\partial_B^2 Y)\,dB\,dB + \cdots \end{aligned} \tag{3.15}$$

What is interesting is that if we only wish to retain terms up to *first* order in dt, we cannot neglect the last term on the right-hand side of (3.15), which is of *second* order in dB. We can certainly see that this is so, since

$$\langle (dB)^2 \rangle \equiv \langle [\, B(t + dt) - B(t)\,]^2 \rangle = \sigma^2 dt \tag{3.16}$$

[property (ii)]. What is *not* obvious is that it is rigorously correct to set $(dB)^2 = \sigma^2 dt$ in equation (3.15), obtaining thereby Itô's rule:

$$dY(t, B) = (\partial_t Y)\,dt + (\partial_B Y)\,dB + \tfrac{1}{2}(\partial_B^2 Y)\sigma^2 dt \tag{3.17}$$

The last term in (3.17)—the *correction term*—would not be present if $B(t)$ was merely a differentiable function.

For example, if we set $Y = \frac{1}{2}B^2$ in equation (3.17), we get

$$d\tfrac{1}{2}B^2(t) = B\,dB + \tfrac{1}{2}\sigma^2 dt \tag{3.18a}$$

We can use this result to integrate equation (3.8):

$$X(t) - X(0) \equiv \int_0^t B(t)\,dB = \tfrac{1}{2}B^2(t) - \tfrac{1}{2}\sigma^2 t \tag{3.18b}$$

[Note that (3.18b) satisfies the Martingale property (3.12a).] The usefulness of Itô's rule in performing stochastic integrals should be clear.

The "derivation" of Itô's rule given above also works in more general cases. Consider N independent Brownian motions $B_1(t),\ldots,B_N(t)$, and M *stochastic differential equations*

$$dX_k = m_K(X_1,\ldots,X_M,t)\,dt + \sum_{l=1}^{N} \sigma_{kl}(X_1,\ldots,X_M,t)\,dB_l, \qquad k=1,\ldots,M \tag{3.19a}$$

Itô's differential of a function $Y(X_1,\ldots,X_M,t)$ is

$$dY = \partial_t Y\,dt + \sum_r \frac{\partial Y}{\partial X_r}\,dX_k + \frac{\sigma^2}{2}\sum_{rsk} \frac{\partial^2 Y}{\partial X_r \partial X_s}\,\sigma_{rk}\sigma_{sk}\,dt \tag{3.19b}$$

Similarly, a heuristic derivation of the differential form of equation (3.7) can be given:

$$\Delta X(t) = \frac{dX}{dt}\,\Delta t + \frac{1}{2}\frac{d^2X}{dt^2}(\Delta t)^2 + \cdots = F\dot{B}\,\Delta t + \frac{1}{2}\frac{d}{dt}F\dot{B}(\Delta t)^2 + \cdots$$

$$= F\dot{B}\,\Delta t + \frac{1}{2}\frac{\partial F}{\partial B}(\dot{B}\,\Delta t)^2 + \frac{1}{2}F\ddot{B}(\Delta t)^2 + \cdots \tag{3.20a}$$

Following the procedure used in deriving Itô's differential rule, we replace $(\dot{B}\,\Delta t)^2 = (\Delta B)^2$ by $\sigma^2\,\Delta t$ and retain this term as it is of first order in Δt, but discard $\ddot{B}(\Delta t)^2$ as being of higher order. The differential form of equation (3.7) is thus

$$dX = F\,dB + \frac{\sigma^2}{2}\frac{\partial F}{\partial B}\,dt \tag{3.20b}$$

and not equation (3.8). The integral of equation (3.20b), when $F = B$, is $X(t) = X(0) + \frac{1}{2}B^2$ according to equation (3.18a) [compare with the integral (3.18b) of equation (3.8)]. This heuristic argument can be rigorously justified, as a rule in what has been called a "Stochastic Calculus" by McShane (1960).

Actually, the kind of differential equation we wish to solve is not really like equation (3.7); it is more like

$$\frac{dX(t)}{dt} = F(X(t))\dot{B}(t) \tag{3.21}$$

In order to integrate (3.21), we must first convert it to a stochastic differential equation. Because of Itô's rule, we do *not* expect that equation

to be $dX = F\,dB$. Indeed, following the argument given above in equations (3.20) we obtain

$$dX = F\,dB + \frac{\sigma^2}{2}\frac{\partial F}{\partial X}F\,dt \tag{3.22}$$

Most importantly for physical applications, it can be rigorously proved that if the (undefined) equation (3.21) is written as the limit of a converging sequence of (defined) equations, where each equation in the sequence has $\dot{B}(t)$ replaced by an approximation to white noise [such as equation (3.4) with a cutoff], the limit of the solutions is the solution $X(t)$ to the stochastic differential equation (3.22). That is, equations (3.21) and (3.22) are equivalent.

Furthermore, it can be shown that if F satisfies certain boundedness and continuity conditions, the solution X of equation (3.22), subject to a known initial condition $X(0)$, is unique, continuous, and is a *Markov process.* A Markov process is a process in which knowledge of the probability distribution of X at any time t suffices for determining the probability distribution of X at all later times. We shall return to this point in a moment.

Equation (3.21) can be generalized to M coupled equations

$$\frac{dX_k}{dt} = G_k(X_1,\ldots,X_M,t) + \sum_{l=1}^{N} F_{kl}(X_1,\ldots,X_M,t)\dot{B}_l(t) \tag{3.23}$$

involving N independent Brownian motions. It is appropriate to regard the solution $\{X_k\}$ of this set of equations as the solution of a set of stochastic differential equations

$$dX_k = G_k\,dt + \sum_{s=1}^{N} F_{ks}\,dB_s + \frac{\sigma^2}{2}\sum_{r=1}^{M}\sum_{s=1}^{N}\frac{\partial F_{ks}}{\partial X_r}F_{rs}\,dt, \qquad k = 1,\ldots,M \tag{3.24}$$

This solution enjoys all the properties cited above for the solution of equation (3.22). Note that the reducing Schrödinger equation has the form (3.23), so we are interested in the solution of equation (3.24).

Because the solution of a stochastic differential equation of the form

$$dX = m(X,t)\,dt + \sigma(X,t)\,dB \tag{3.25}$$

is a Markov process, it can be completely characterized in terms of the conditional probability $\rho(x,t|x_0,t_0)\,dx$ for X to lie between x and $x+dx$ at time t, given that X took on the value x_0 at time t_0. It may be shown that ρ

is in turn completely characterized by the mean and variance of $dX(t_0)$, which can be directly calculated from (3.25):

$$\langle dX(t_0)\rangle = m(x_0,t_0)\,dt \tag{3.26a}$$

$$\langle\{dX(t_0)-\langle dX(t_0)\rangle\}^2\rangle = \sigma^2\sigma^2(x_0,t_0)\,dt \tag{3.26b}$$

[recall that $X(t_0)$ and $dX(t_0)$ are statistically independent]. The characterization is in terms of two equations that ρ must satisfy:

$$-\frac{\partial\rho}{\partial t_0} = m(x_0,t_0)\frac{\partial\rho}{\partial x_0} + \frac{1}{2}\sigma^2\sigma^2(x_0,t_0)\frac{\partial^2\rho}{\partial x_0} \tag{3.27a}$$

$$+\frac{\partial\rho}{\partial t} = -\frac{\partial}{\partial x}m(x,t)\rho + \frac{\sigma^2}{2}\frac{\partial^2}{\partial x^2}\sigma^2(x,t)\rho \tag{3.27b}$$

These are the *backward diffusion equation*, and the *forward diffusion*—or *Fokker–Planck*—equation, respectively. ρ can be determined by solving either one, subject to the initial condition $\rho(x,t_0|x_0,t_0)=\delta(x-x_0)$.

To summarize: one can start with a differential equation linear in white noise such as equation (3.21), use Itô's rule to construct the associated stochastic differential equation (3.22) for the solution, which has the form (3.25), and examine the statistical behavior of the solution X by investigating the solution ρ of the Fokker–Planck equation (3.27b).

We shall need the generalization of (3.26), (3.27b) to the *set* of stochastic differential equations (3.19a):

$$\langle dX_k\rangle = m_k\,dt \tag{3.28a}$$

$$\langle\{dX_k-\langle dX_k\rangle\}\{dX_l-\langle dX_l\rangle\}\rangle = \sigma^2\,dt\sum_n \sigma_{km}\sigma_{lm} \equiv \sigma^2\,dt(\sigma\tilde{\sigma})_{kl} \tag{3.28b}$$

$$\frac{\partial\rho}{\partial t} = -\sum_k \frac{\partial}{\partial x_k}m_k\rho + \frac{\sigma^2}{2}\sum_{kl}\frac{\partial^2}{\partial x_k\partial x_l}(\sigma\tilde{\sigma})_{kl}\rho \tag{3.29}$$

4. REDUCING EQUATION

We now begin our investigation of the conditions under which equation (2.8) becomes a reducing Schrödinger equation. It is convenient to rewrite equation (2.8) in terms of the amplitudes $a_n \equiv \langle\phi_n(t)|\psi(t)\rangle$:

$$i\hbar\frac{da_n}{dt} = \hbar\omega_n a_n + \langle\phi_n|(H-H_0)|\psi\rangle + \sum_m A_{mn}\dot{B}_{mn}F(a_m,a_n) \tag{4.1}$$

To fix ideas, imagine that we are describing the interaction of a microscopic system with an apparatus. The states $|\phi_n\rangle$ describe the combined microscopic system plus apparatus. Initially, the states $|\phi_n\rangle$ that have nonzero amplitudes a_n are assumed to be macroscopically indistinguishable. By hypothesis, the matrix elements $A_{mn}=\langle\phi_m|A|\phi_n\rangle$ connecting these states are negligibly small, so equation (4.1) is equivalent to the usual Schrödinger equation. After the interaction takes place, on the other hand, it is assumed that $\langle\phi_n|H|\phi_m\rangle\simeq\langle\phi_n|H_0|\phi_m\rangle$, i.e., the Hamiltonian H no longer effects transitions between the states $|\phi_n\rangle$, so equation (4.1) effectively becomes

$$i\hbar\frac{da_n}{dt}=\hbar\omega_n a_n+\sum_m A_{mn}\dot{B}_{mn}F(a_m,a_n) \tag{4.2}$$

The states $|\phi_n\rangle$ for which a_n is nonzero are now macroscopically distinguishable, and the matrix elements A_{mn} are large (for simplicity, we shall assume $A_{nn}\equiv 0$ is always true).

Equation (4.2) will be our starting point: we will examine its solutions for $t\geqslant 0$, with initial conditions $a_n(0)$ assumed given. Of course, the transition from Schrödinger's equation to equation (4.2) is not abrupt. There is a period of time during which both interaction terms on the right-hand side of equation (4.1) are of comparable magnitude. It is possible that there may be experiments (Papaliolios, 1967)[2] that can measure the interference effects between these two terms and thereby test the correctness of this theory. The methods of this paper can be used to investigate these interference effects, but we shall not do so here.

Our first step is to write the stochastic differential equation that is equivalent to equation (4.2). Equation (4.2) differs from the comparable set of white noise differential equations (3.23) of the previous section in that the variables are complex, and the white noise is complex, i.e., $\dot{B}_{mn}=\dot{B}_{Rmn}+i\dot{B}_{IMn}$, where $\dot{B}_{Rmn}$ and $\dot{B}_{Imn}$ are independent real white noise functions ($\dot{B}_{Rmn}=\dot{B}_{Rnm}, \dot{B}_{Imn}=-\dot{B}_{Inm}$). One can of course take the real and imaginary parts of equation (4.2) and treat them as in Section 3. A simpler approach is to treat a_n and a_n^* as independent variables, and to be careful to apply Itô's rule using

$$dB_{mn}dB_{rs}=\delta_{ms}\delta_{nr}2\sigma^2 dt \tag{4.3}$$

Using the procedure of the previous section, we find that the

[2]Also, see the discussion in Belinfante (1973).

stochastic differential equations equivalent to (4.2) are

$$da_n = -i\omega_n a_n\,dt + \frac{1}{i\hbar}\sum_m A_{mn}F_{mn}\,dB_{mn} + \frac{\sigma^2}{\hbar^2}\,dt\sum_m |A_{mn}|^2\left(\frac{\partial F_{mn}}{\partial a_n^*}F_{mn}^* - \frac{\partial F_{mn}}{\partial a_m}F_{nm}\right) \tag{4.4}$$

and the complex conjugate equations. In deriving equations (4.4) we have defined

$$F_{mn} \equiv F(a_m, a_n, a_m^*, a_n^*) \tag{4.5}$$

making explicit the independent variables upon which F depends.

In order to obtain the Fokker–Planck equation describing the statistical distribution of the solutions of equations (4.4), we need the means and nonvanishing variances. These are calculated from equations (4.4) to be

$$\langle da_n\rangle = -i\omega_n a_n\,dt + \frac{\sigma^2}{\hbar^2}\,dt\sum_m |A_{mn}|^2\left[\frac{\partial F_{mn}}{\partial a_n^*}F_{mn}^* - \frac{\partial F_{mn}}{\partial a_m}F_{nm}\right] \tag{4.6a}$$

$$\langle[da_n^* - \langle da_n^*\rangle][da_n - \langle da_n\rangle]\rangle = \frac{2\sigma^2}{\hbar^2}\,dt\sum_m |A_{mn}F_{mn}|^2 \tag{4.6b}$$

$$\langle[da_m - \langle da_m\rangle][da_n - \langle da_n\rangle]\rangle = -\frac{2\sigma^2}{\hbar^2}\,dt|A_{mn}|^2F_{mn}F_{nm}, \qquad m\neq n \tag{4.6c}$$

and the complex conjugates of equations (4.6a), (4.6c). The Fokker–Planck equation is

$$\frac{\partial\rho}{\partial t} = -\sum_n\left[\frac{\partial}{\partial a_n}\frac{\langle da_n\rangle}{dt}\rho + \frac{\partial}{\partial a_n^*}\frac{\langle da_n^*\rangle}{dt}\rho\right] + \frac{\sigma^2}{\hbar^2}\sum_{n,m} 2\frac{\partial^2}{\partial a_n^*\partial a_n}|A_{mn}F_{mn}|^2\rho - \frac{\sigma^2}{\hbar^2}\sum_{n,m}\left[\frac{\partial^2}{\partial a_n\partial a_m}|A_{mn}|^2F_{mn}F_{nm}\rho + \frac{\partial^2}{\partial a_n^*\partial a_m^*}|A_{mn}|^2F_{mn}^*F_{nm}^*\rho\right] \tag{4.7}$$

Before seeking a reducing Schrödinger equation, it is interesting to apply these results to the ordinary Schrödinger equation. If we set $F = a_m$ in equation (4.2) we have the ordinary Schrödinger equation with a

random Hamiltonian. The Fokker–Planck equation (4.7) is then

$$\frac{\partial \rho}{\partial t} = i \sum_n \omega_n \left(\frac{\partial}{\partial a_n} a_n \rho - \frac{\partial}{\partial a_n^*} a_n^* \rho \right)$$
$$+ \frac{\sigma^2}{\hbar^2} \sum_{n,m} |A_{mn}|^2 \left[\frac{\partial}{\partial a_n} a_n \rho + \frac{\partial}{\partial a_n^*} a_n^* \rho + 2 \frac{\partial^2}{\partial a_n \partial a_n^*} a_m a_m^* \rho - \frac{\partial^2}{\partial a_m \partial a_n} a_m a_n \rho \right.$$
$$\left. - \frac{\partial^2}{\partial a_m^* \partial a_n^*} a_m^* a_n^* \rho \right] \tag{4.8}$$

Because we are mostly interested in the behavior of the magnitudes $|a_n|$, it is convenient to change to real variables x_n, θ_n defined by

$$a_n = (x_n)^{1/2} e^{i\theta_n} \tag{4.9}$$

and the complex conjugate equation. In terms of these variables, equation (4.8) becomes

$$\frac{\partial \rho}{\partial t} = \sum_n \omega_n \frac{\partial \rho}{\partial \theta_n} + \frac{\sigma^2}{\hbar^2} \sum_{n,m} |A_{mn}|^2 \left[\left(\frac{\partial}{\partial x_n} - \frac{\partial}{\partial x_m} \right) x_n x_m \left(\frac{\partial}{\partial x_n} - \frac{\partial}{\partial x_m} \right) \rho \right.$$
$$\left. + \frac{1}{4} \frac{\partial^2}{\partial \theta_n^2} \left(\frac{x_m}{x_n} \rho \right) + \frac{1}{4} \frac{\partial^2}{\partial \theta_m^2} \left(\frac{x_n}{x_m} \rho \right) + \frac{1}{2} \frac{\partial^2 \rho}{\partial \theta_n \partial \theta_m} \right] \tag{4.10}$$

Equation (4.10) provides all the necessary information about the statistical behavior of the solutions of the Schrödinger equation. For example, we may find the behavior of $\langle x_k \rangle$ by multiplying equation (4.10) by x_k and integrating over all x_n, θ_n:

$$\frac{d}{dt} \langle x_k \rangle = \frac{\sigma^2}{\hbar^2} \sum_m |A_{mk}|^2 \{ \langle x_m \rangle - \langle x_k \rangle \} \tag{4.11}$$

Equation (4.11) has the form of Pauli's (1926) "master equation" based upon the second-order perturbation theory "golden rule." It differs in that (1) equation (4.11) is exact, (2) the expectations $\langle x_n \rangle$ appear in equation (4.11) (not the probabilities x_n themselves), and (3) there is no energy-conserving δ function in equation (4.11) (because the white-noise Hamiltonian contains all frequencies, and so effects transitions between states of any energy). It can be seen from equation (4.11) that asymptotically, where $\langle \dot{x}_n \rangle = 0$, all $\langle x_n \rangle$ become equal.

A diffusion equation for the amplitudes x_n alone can be obtained by integrating equation (4.10) over all angles:

$$\frac{\partial \rho}{\partial t} = \frac{\sigma^2}{h^2} \sum_{n,m} |A_{mn}|^2 \left(\frac{\partial}{\partial x_n} - \frac{\partial}{\partial x_m} \right) x_n x_m \left(\frac{\partial}{\partial x_n} - \frac{\partial}{\partial x_m} \right) \rho \tag{4.12}$$

(we have written $\int d\boldsymbol{\theta} \rho \rightarrow \rho$). Equation (4.12) was also arrived at in I as a consequence of the ordinary Schrödinger equation with a random initial phase approximation. It was shown there that equation (4.12) implies that the probability density ρ asymptotically uniformly fills all the available "phase space" ($\sum x_n = 1, 0 \leqslant x_n \leqslant 1$).

We now leave this example of the use of this formalism and return to the main problem: how to choose the function F so that properties 0, 1, and 2 will be satisfied, and equation (4.2) will be a reducing equation.

Property 0 will be achieved by the vanishing of

$$i\hbar \frac{d}{dt} \sum_n a_n^* a_n = \sum_{m,n} A_{mn} \dot{B}_{mn} (a_n^* F_{mn} - a_m F_{nm}^*) \tag{4.13}$$

Because of the independence of the $\dot{B}_{mn}$, equation (4.13) will be satisfied if and only if

$$a_n^* F_{mn} = a_m F_{nm}^* \tag{4.14}$$

In order to impose *properties 1 and 2* in the form of equations (1.3), we multiply the Fokker–Planck equation (4.7) by $a_n^* a_n$ and integrate over all x_k, θ_k, obtaining

$$\frac{d\langle x_n \rangle}{dt} = \left\langle a_n^* \frac{\langle da_n \rangle}{dt} + a_n \frac{\langle da_n^* \rangle}{dt} + \frac{2\sigma^2}{h^2} \sum_m |A_{mn}|^2 |F_{mn}|^2 \right\rangle \tag{4.15}$$

[$d\langle a_n \rangle / dt$ is given by equation (4.6a)]. To make $d\langle x_n \rangle / dt$ vanish identically, we require the quantity within the brackets on the right-hand side of equation (4.15) multiplying each $|A_{mn}|^2$ to vanish:

$$a_n^* \left[\frac{\partial F_{mn}}{\partial a_n^*} F_{mn}^* - \frac{\partial F_{mn}}{\partial a_m} F_{nm} \right] + \text{c.c.} + 2|F_{mn}|^2 = 0 \tag{4.16}$$

(c.c. is the complex conjugate of the first term). This equation, to which we will shortly return, is solved in Appendix B.

Equation (1.3b) can be implemented by multiplying the Fokker–Planck equation (4.7) by $a_n^* a_n a_m^* a_m$ and integrating over all x_k, θ_k. If we use

equation (4.14) and equation (4.16) to simplify the result, we readily obtain

$$\frac{d}{dt}\langle x_n x_m\rangle = \frac{-4\sigma^2}{\hbar^2}\langle x_m|F_{nm}|^2\rangle = \frac{-4\sigma^2}{\hbar^2}\langle x_n|F_{mn}|^2\rangle \tag{4.17}$$

Equation (4.17) implies that asymptotically

$$x_n|F_{mn}|^2 \underset{t\to\infty}{\to} 0 \tag{4.18}$$

This is because $\langle x_n x_m\rangle$ is a positive semidefinite function, and it is monotonically decreasing by equation (4.17). Therefore, it must asymptotically achieve a lower bound, which means that the expectation value on the right-hand side of equation (4.17) must asymptotically vanish. But as $x_n|F_{mn}|^2$ is also positive semidefinite, its expectation value can only vanish if it asymptotically vanishes for each sample.

Now if F is such that

$$x_n|F_{mn}|^2 = 0 \qquad \text{implies } x_n x_m = 0 \tag{4.19}$$

(for example, if $|F_{mn}|^2 = x_n^{r-1}x_m^r, r>0$), then equation (1.3b) is obtained.

The most general form of F satisfying equations (4.14) and (4.16) is found in Appendix B [equation (B.8)]. It is argued there that the simplest (but by no means the only) choice that also satisfies equation (4.19) is

$$F_{mn} \equiv a_m^* a_n / a_n^* \tag{4.20}$$

It is easy to see that (4.20) satisfies (4.14), (4.16), and (4.19), and so equation (4.2) in the form

$$i\hbar\frac{da_n}{dt} = \hbar\omega_n a_n + \sum_M A_{mn}\dot{B}_{mn}a_m^*\frac{a_n}{a_n^*} \tag{4.21}$$

entails properties 0, 1, 2 and is a reducing equation.

In this specific case, the stochastic differential equations equivalent to equations (4.21) are

$$da_n = -i\omega_n a_n + \frac{1}{i\hbar}\sum_m A_{mn}\frac{a_m a_n}{a_n^*}\,dB_{mn} - \frac{\sigma^2 dt}{\hbar^2}\sum_m |A_{mn}|^2\frac{|a_m|^2}{a_n^*} \tag{4.22}$$

according to equation (4.4). The associated Fokker–Planck equation is, by

equations (4.6), (4.7)

$$\frac{\partial\rho}{\partial t} = i\sum_n \omega_n\left(\frac{\partial}{\partial a_n}a_n\rho - \frac{\partial}{\partial a_n^*}a_n^*\rho\right)$$
$$+\frac{\sigma^2}{\hbar^2}\sum_{n,m}|A_{mn}|^2\left[\frac{\partial}{\partial a_n}\frac{|a_m|^2}{a_n^*}\rho + \frac{\partial}{\partial a_n^*}\frac{|a_m|^2}{a_n}\rho + 2\frac{\partial^2}{\partial a_n\partial a_n^*}a_m a_m^*\rho\right.$$
$$\left. -\frac{\partial^2}{\partial a_m\partial a_n}a_m a_n\rho - \frac{\partial^2}{\partial a_m^*\partial a_n^*}a_m^* a_n^*\rho\right] \tag{4.23}$$

If we change to variables x_n, θ_n according to equation (4.9), then equation (4.23) becomes

$$\frac{\partial\rho}{\partial t} = \sum_n \omega_n\frac{\partial\rho}{\partial\theta_n} + \frac{\sigma^2}{\hbar^2}\sum_{n,m}|A_{mn}|^2\left[\left(\frac{\partial}{\partial x_n} - \frac{\partial}{\partial x_m}\right)^2(x_n x_m\rho)\right.$$
$$\left. +\frac{1}{4}\frac{\partial^2}{\partial\theta_n^2}\left(\frac{x_m}{x_n}\rho\right) + \frac{1}{4}\frac{\partial^2}{\partial\theta_m^2}\left(\frac{x_n}{x_m}\rho\right) + \frac{1}{2}\frac{\partial^2\rho}{\partial\theta_n\partial\theta_m}\right] \tag{4.24}$$

This should be compared with the Fokker–Planck equation (4.10) for the ordinary Schrödinger equation. (Also compare equation 4.25 below with equation 4.12).

A diffusion equation for the amplitudes x_n alone can be found by integrating equation (4.24) over all angles:

$$\frac{\partial\rho}{\partial t} = \frac{\sigma^2}{\hbar^2}\sum_{n,m}|A_{mn}|^2\left(\frac{\partial}{\partial x_n} - \frac{\partial}{\partial x_m}\right)^2 x_n x_m\rho \tag{4.25}$$

(we have written $\int d\boldsymbol{\theta}\rho\rightarrow\rho$). This equation, introduced in I, was discussed in Section 2 [equation (2.6)]. It is easily seen to imply equations (2.7a)–(2.7c), which embody properties 0, 1, and 2.

5. DEFINING $|\phi_n\rangle$ AND H_0

In this paper and in I, mechanisms have been presented that might be used in a complete theory that explains why events occur. What is perhaps surprising is that it is fairly easy to construct a reducing equation just by adding a term to the Schrödinger equation, that the term need not be terribly complicated, and that there is a reasonably large class of terms that will do the job.

However, it should be emphasized that so far this is not a complete theory. There are still unspecified elements: the observable states $|\phi_n\rangle$, their time development Hamiltonian H_0, and the interaction operator A. Let us discuss these.

In I we left the definition of $|\phi_n\rangle$ and H_0 open, but here we will venture a response to the problem, following closely the work of Kübler and Zeh (1973). The point is that a natural basis exists if we split the physical system under discussion into two parts, a "small" part and a "large" part. For example, we may consider the small part to be a single microscopic system, and the large part to be an apparatus making a measurement on this system. Or, we may consider the small part to be a microscopic system plus apparatus, and the large part to be the "environment" surrounding the apparatus.

The state vector together with this split define a basis at every instant of time because it is always possible to expand the state vector $|\psi\rangle$ in the "Schmidt canonical form"

$$|\psi\rangle = \sum_n a_n |\sigma_n\rangle |\lambda_n\rangle \tag{5.1}$$

where $\{|\sigma_n\rangle\}$ is an orthonormal basis for the small part, and $\{|\lambda_n\rangle\}$ is an orthonormal basis for the large part. This expansion is "almost" unique (the meaning of that will be clarified in a moment). Of course, what makes equation (5.1) so strict a requirement is the absence of a double sum over the complete basis $|\sigma_n\rangle|\lambda_m\rangle$. If one describes the measurement as is customary, by the density matrix

$$\rho_\sigma \equiv \mathrm{Tr}_\lambda |\psi\rangle\langle\psi| = \sum_n |\sigma_n\rangle |a_n|^2 \langle\sigma_n| \tag{5.2}$$

the benefit of the form (5.1) is apparent: the basis $|\sigma_n\rangle$ is the one in which the density matrix is diagonal, and arguments have been given (Jauch, 1968; Zeh, 1970, 1971) for the two examples mentioned above that these are actually the states observed in nature. We suggest that our physical states be identified as

$$|\phi_n\rangle = |\sigma_n\rangle |\lambda_n\rangle \tag{5.3}$$

The decomposition (5.1) is unique, up to choices of phase factors multiplying $|\sigma_n\rangle, |\lambda_n\rangle$, if the magnitudes a_n are all different. We will fix the phase factors (up to an initial phase factor) by requiring

$$\langle\sigma_n|d|\sigma_n\rangle / dt = \langle\lambda_n|d|\lambda_n\rangle / dt = 0 \tag{5.4}$$

If two magnitudes are equal, say $|a_n| = |a_m|$, but the nth and mth states are macroscopically indistinguishable, it does not matter how we choose the two basic vectors to span the n-m subspace. If the nth and mth states are superpositions of two macroscopically distinguishable states, we choose the macroscopically distinguishable states as basis vectors.

There is yet another ambiguous aspect: how do we decide where the small part of the physical system under consideration ends and the large part begins? Our expectation is that the predictions of the theory are relatively insensitive to a reasonable location of this division.

This method of choosing a basis automatically determines the time evolution of the states $|\phi_n\rangle$, i.e., determines the Hamiltonian H_0. To see this, we write the reduction equation (2.5) as

$$i\hbar \frac{d}{dt}|\psi\rangle = i\hbar \sum_n \{\dot{a}_n|\sigma_n\rangle|\lambda_n\rangle + a_n|\dot{\sigma}_n\rangle|\lambda_n\rangle + a_n|\sigma_n\rangle|\dot{\lambda}_n\rangle\}$$

$$= H|\psi\rangle + \sum_{m,n} |\sigma_n\rangle|\lambda_n\rangle \frac{a_n}{a_n^*} a_m^* A_{mn}^{mn} \qquad \left(A_{kl}^{ij} \equiv \langle\lambda_l|\langle\sigma_k|A|\sigma_1\rangle|\lambda_j\rangle\right) \tag{5.5}$$

Upon taking the scalar product of equation (5.5) with $\langle\sigma_n|\langle\lambda_n|$ we obtain the reduction equation for the amplitudes:

$$i\hbar\dot{a}_n = \bar{H}_{0nn}a_n + \bar{H}^{0nn}a_n + \sum_m \left\{ H_{Inm}^{nm}a_m + \frac{a_n}{a_n^*} a_m^* A_{mn}^{mn} \right\} \tag{5.6}$$

In equation (5.6) we have written $H = \bar{H}_0 \times 1 + 1 \times \bar{H}^0 + H_I$, where $\bar{H}_0$ is the part of the Hamiltonian operating solely on the $\{|\sigma_n\rangle\}$ subspace, $\bar{H}^0$ acts only on the $\{|\lambda_n\rangle\}$ subspace, and H_I is the interaction part of the Hamiltonian. Connection with a previous notation is made by the identification

$$\hbar\omega_n \equiv \bar{H}_{0nn} + \bar{H}^{0nn} \tag{5.7}$$

The scalar product of equation (5.5) with $\langle\sigma_n|\langle\lambda_m|$ $(m \neq n)$ yields

$$i\hbar a_m\langle\sigma_m|\dot{\sigma}_n\rangle + iha_n\langle\lambda_m|\dot{\lambda}_n\rangle = \sum_s H_{ns}^{ms}a_s \tag{5.8}$$

Note that there is no direct dependence on A in equations (5.8). Equations

(5.8) are completely solvable, and the solution is (Kübler and Zeh, 1973)

$$i|\dot{\sigma}_n\rangle = \sum_{m,s}{}' |\sigma_m\rangle \left[H^{ns}_{ms} a_s a_n^* - H^{sm}_{sn} a_s^* a_m \right] / \left[|a_n|^2 - |a_m^2| \right] \tag{5.9a}$$

$$i|\dot{\lambda}_n\rangle = \sum_{m,s}{}' |\lambda_m\rangle \left[H^{ms}_{ns} a_s a_n^* - H^{sn}_{sm} a_s^* a_m \right] / \left[|a_n|^2 - |a_m|^2 \right] \tag{5.9b}$$

[The prime in the sum denotes that $m \neq n$: also, the right-hand sides of equations (5.9) do not become infinite if $|a_n| = |a_m|$, $m \neq n$, as the dynamics must be such that the numerator vanishes too.] We find $d|\sigma_n\rangle|\lambda_n\rangle / dt$ from equations (5.9) to be

$$i\hbar \frac{d}{dt}|\phi_n\rangle = \left(\overline{H}_0 + \overline{H}^0 - \hbar\omega_n + \tilde{H}_0\right)|\phi_n\rangle$$

$$\tilde{H}_0|\phi_n\rangle \equiv |\lambda_n\rangle \sum_{ms}{}' |\sigma_m\rangle \frac{\left\{ H^{ns}_{Ims} a_s a_n^* - H^{sm}_{Isn} a_s^* a_m \right\}}{|a_n|^2 - |a_m|^2}$$

$$+ |\sigma_n\rangle \sum_{ms}{}' |\lambda_m\rangle \frac{\left\{ H^{ms}_{Ins} a_s a_n^* - H^{sn}_{Ism} a_s^* a_m \right\}}{|a_n|^2 - |a_m|^2} \tag{5.10}$$

From equations (2.2) and (5.10) we identify

$$H_0 \equiv \overline{H}_0 + \overline{H}^0 + \tilde{H}_0 \tag{5.11}$$

(note that $\langle \phi_m | H_0 | \phi_n \rangle = \hbar\omega_n \delta_{nm}$) as the Hamiltonian governing the time evolution of the state $|\phi_n\rangle$.

Before closing this section, we wish to make two more points.

First, in the context of this discussion, we can see the random initial phase and the random interaction mechanisms as two extremes. If the matrix elements A^{mn}_{mn} in equation (5.6) are essentially constant, we invoke the first mechanism, while, if the matrix elements fluctuate rapidly (perhaps due to their dependence on the large part state vectors, or perhaps simply due to a fundamental randomness in nature), we invoke the second mechanism.

Last, how might we apply this method of choosing the observable basis to the physical system consisting of the whole universe? It would be surprising if the universe could be split into two pieces so that in the Schmidt decomposition (5.1) both $|\sigma_n\rangle$ and $|\lambda_n\rangle$ are almost always observable states, i.e., are not superpositions of macroscopically different states. Perhaps it is possible to split the universe into many smaller pieces in such

a way that, in the Schmidt-like decomposition

$$|\psi\rangle = \sum_n a_n |\sigma_{1n}\rangle |\sigma_{2n}\rangle |\sigma_{3n}\rangle \cdots \tag{5.12}$$

each $|\sigma_{in}\rangle$ corresponds to an observable state. This raises questions about the uniqueness of the decomposition (5.12), about the size of the "cells" into which the universe must be divided, about the effectiveness of the Hamiltonian in coupling cells, etc.

6. CONCLUDING REMARKS

It is still an open question as to what the operator A should be that governs the reduction process. We have stated constraints on it: its matrix elements should be small between macroscopically distinguishable states. How can this be achieved? After all, states can differ macroscopically in many ways: in energy, momentum, angular momentum, position, etc.

We conjecture that, if $|\psi\rangle$ is initially a superposition of states that are "close" in energy, it is sufficient that the magnitude of matrix elements of A depends only upon the macroscopic distinguishability of appropriate position variables characterizing the states. The thought here is that, if two states close in energy should differ by a macroscopic variable that is not position, the normal dynamics of the states according to the Schrödinger equation will soon make them differ in position as well. After that, reduction to one of the states will proceed because A_{mn} will be large.

It is appropriate to make the following point here. In the previous sections of this paper we have, for simplicity, assumed that the matrix elements A_{mn} are *all* large simultaneously, so that the reduction will be to precisely one of the states $|\phi_n\rangle$ in the superposition (1.1). However, if A behaves as outlined above, the matrix elements A_{mn} are only large between macroscopically distinguishable states, and the reduction proceeds to a *family* of macroscopically indistinguishable states.

The dynamics of this process is as follows. For definiteness, suppose that

$$|\psi\rangle = \sum_n a_n |\phi_n\rangle + \sum_m b_m |\chi_m\rangle \tag{6.1}$$

where $\{|\phi_n\rangle\}$ are macroscopically indistinguishable, and $\{|\chi_m\rangle\}$ are macroscopically indistinguishable, but each $|\phi_n\rangle$ differs macroscopically from each $|\chi_m\rangle$. This might be considered as a model for an experiment with just two outcomes, with the set $\{|\phi_n\rangle\}$ representing various microscopic states of a single "pointer" position, and similarly for the set $\{|\chi_n\rangle\}$. Let us

define $x_n \equiv |a_n|^2$, $y_m \equiv |b_m|^2$, $x \equiv \Sigma x_n$, $y \equiv \Sigma y_m$ $(x+y=1)$. What happens may be likened to two rows of soldiers shooting at each other. It is easily shown by taking first and second moments of equation (4.25).

$$\frac{\partial \rho}{\partial t} = \frac{2\sigma^2}{\hbar^2} \sum_{n,m} |A_{mn}|^2 \left[\frac{\partial}{\partial x_n} - \frac{\partial}{\partial y_m} \right]^2 x_n y_m \rho \tag{6.2}$$

that

$$\langle x(t) \rangle = x(0), \qquad \langle y(t) \rangle = y(0), \qquad \langle x(t) y(t) \rangle \underset{t \to \infty}{\to} 0 \tag{6.3}$$

which implies that either the set $\{x_n\}$ asymptotically vanishes and $y(\infty)=1$ [with probability $y(0)$], or the set $\{y_m\}$ asymptotically vanishes and $x(\infty)=1$ [with probability $x(0)$].

It is probably worthwhile to try to construct an operator A with the desired characteristics, and to test its behavior in calculations, especially those involving interference between the Hamiltonian and the reduction term (Papaliolios, 1967). But ultimately, a theory such as this needs to be legitimized by being a consequence of a larger theory that has more ties to established physics. In what areas might such a theory arise? Outside of the present application, there seems to be at present no groundswell of need for a nonlinear quantum theory (but see Bialynicki-Birula and Mycielski, 1976). Perhaps we might look to the operator A for a suggestion. It is a nonlocal long-range interaction between a system and itself. This carries the connotation of relevance to self-energy considerations, and perhaps gravitational theory. Indeed, it is an attractive thought that the juncture between general relativity, which describes events but does not describe microscopic behavior, and quantum theory, which describes microscopic behavior but does not describe events, might be an appropriate place to look.

APPENDIX A

In this appendix we give a more familiar proof than those given in I that equation (2.5) is a reducing equation in second-order perturbation theory. As in Section 4, we express equation (2.5) in terms of the amplitudes $a_n \equiv \langle \phi_n(t) | \psi(t) \rangle$:

$$i\hbar \frac{da_n}{dt} = \hbar\omega_n + \langle \phi_n | (H - H_0) | \psi \rangle + \sum_m A_{mn} \frac{a_m^* a_n}{a_n^*} \tag{A.1}$$

and consider solving equation (A.1) for $t \geqslant 0$, when $\langle \phi_n | (H - H_0) | \psi \rangle \simeq 0$.

We first integrate equation (A.1) using standard second-order perturbation theory. With the substitution $b_n = a_n e^{-i\omega_n t}$, the integral of equation (A.1) is

$$b_n(t) = b_n(0) - i\hbar^{-1}\int_0^t \sum_m A_{mn} b_m^*(t_1) b_n(t_1) b_n^{*-1}(t_1) e^{i\omega_{nm}t_1} dt_1 \quad \text{(A.2)}$$

($\omega_{nm} \equiv \omega_n - \omega_m$). We assume that the matrix elements A_{mn} can be taken out of the time integrals because they vary slowly enough over the time interval $\simeq \hbar^2/\Sigma_m |A_{mn}|^2 \delta(\omega_{mn})$. When equation (A.2) is iterated to second order we obtain

$$\begin{aligned} b_n(t) = b_n(0) &- i\hbar^{-1}\sum_m A_{mn} b_m^*(0) b_n(0) b_n^{*-1}(0)\zeta_{nm}(t) \\ &- \hbar^{-2}\sum_{m,r} A_{mn}A_{rm}^* b_n(0) b_r(0) b_m^*(0) b_m^{-1}(0) b_n^{*-1}(0)\zeta_{nmrm}(t) \\ &- \hbar^{-2}\sum_{m,r} A_{mn}A_{rn}^* b_n(0) b_r^*(0) b_m^*(0) b_n^{*-2}(0)\zeta_{nmnr}(t) \\ &- \hbar^{-2}\sum_{m,r} A_{mn}A_{rn}^* b_r(0) b_m^*(0) b_n^{*-1}(0)\zeta_{nmrn}(t) \end{aligned} \quad \text{(A.3)}$$

$$\zeta_{nm}(t) \equiv \int_0^t e^{i\omega_{nm}t_1} dt_1 \quad \text{(A.4a)}$$

$$\zeta_{nmkl}(t) \equiv \int_0^t e^{i\omega_{nm}t_1}\zeta_{kl}(t_1)\, dt_1 \quad \text{(A.4b)}$$

It is assumed that, to an ensemble of experiments described by equation (A.1), at time $t=0$ there corresponds an ensemble of initial conditions that have identical values of $x_n(0)$, but completely random $\theta_n(0)$ [$b_n(0) = a_n(0) = [x_n(0)]^{1/2} \exp i\theta_n(0)$]. If we denote by $\langle\ \rangle$ the ensemble average over all phase angles, we can make use of such relations as

$$\langle b_n^*(0) b_m(0)\rangle = \delta_{nm} x_n(0) \quad \text{(A.5a)}$$

$$\langle b_n(0) b_r(0) b_m^*(0) b_m^{-1}(0)\rangle = \delta_{nm}\delta_{rm} x_m^2(0) \quad \text{(A.5b)}$$

etc. It is then straightforward to calculate to second order

$$\langle x_n(t)\rangle = x_n(0) \quad \text{(A.6a)}$$

$$\langle\{x_n(t) - x_n(0)\}\{x_m(t) - x_m(0)\}\rangle = -2\hbar^{-2}|A_{mn}|^2 x_m(0) x_n(0) S_{mn}(t), \quad n \neq m \quad \text{(A.6b)}$$

$$\langle\{x_n(t)-x_n(0)\}^2\rangle = 2\hbar^{-2}\sum_m |A_{mn}|^2 x_m(0)x_n(0)S_{mn}(t) \tag{A.6c}$$

$$S_{mn}(t) \equiv \frac{\sin^2(\omega_{mn}t/2)}{(\omega_{mn}t/2)^2} = t2\pi\delta(\omega_{mn}) \tag{A.7}$$

(we have used $A_{nn}=0$). {This may be compared with a similar calculation for the ordinary Schrödinger equation $i\hbar\dot{a}_n = \hbar\omega_n a_n + \sum_m A_{nm}a_m$: equation (A.6a) is replaced by

$$\langle x_n(t)\rangle = x_n(0) + \hbar^{-2}\sum_m |A_{mn}|^2[\,x_m(0)-x_n(0)\,]S_{mn}(t)$$

while equations (A.6b) and (A.6c) are unchanged.}

We make the usual assumptions involved in second-order time-dependent perturbation theory, such as many states closely spaced in energy, and $|A_{mn}|^2$ varying slowly with ω_n, and

$$0 < t \ll \frac{\hbar^2}{\Sigma_m |A_{mn}|^2 \delta(\omega_{mn})} \tag{A.8}$$

Because of the restriction (A.8) on the time interval over which the second-order perturbation term dominates the solution, we are faced with the problem of extending the solution past this time interval.

If we do the second-order perturbation calculation above for the interval $(t+T, T)$ instead of the interval $(t, 0)$, and *assume* that the phases of $b_m(T)$ are randomly distributed, the results are identical to equations (A.6) (with the obvious replacements of t by $t+T$, and 0 by T). But is this assumption of random phases at a time greater than zero justified? It is in this case. We obtained equations (A.6) from equations (A.3) by using equations (A.5), but actually we only needed equation (A.5a) to be true to first order and equations like (A.5b) to be true to zeroth order for the argument to go through. In deriving equations (A.6) for the interval $(2t, t)$ we need

$$\langle b_n^*(t)b_m(t)\rangle = x_n(t)\delta_{nm} + O(|A|^2) \tag{A.9a}$$

$$\langle b_n(t)b_r(t)b_m^*(t)b_m^{-1}(t)\rangle = \delta_{nm}\delta_{rm}x_m^2(t) + O(|A|), \qquad \text{etc.} \tag{A.9b}$$

Equation (A.9b) is obviously correct, since $b_n(t) = b_n(0)$ to zeroth order. It is straightforward to calculate from equation (A.3) that equation (A.9a)

also holds—in fact, to second order. Thus the random phase assumption at $t=0$ implies the phases are also random at a later time $t>0$.[3]

One is therefore justified in extending equations (A.6) over successive intervals $(t,0)$, $(2t,t)$, $(3t,2t)$, etc. We may now treat the solutions of the reduction equation as a Markov process whose means and variances are given by equations (A.6). We can immediately write down the associated Fokker–Planck equation (3.29) using equations (A.6):

$$\frac{\partial\rho}{\partial t}=\frac{\pi}{\hbar^2}\sum_{m,n}|A_{mn}|^2\delta(\omega_{mn})\left(\frac{\partial}{\partial x_n}-\frac{\partial}{\partial x_m}\right)^2 x_n x_m\rho \tag{A.10}$$

This is equation (2.6), which guarantees reduction behavior, as was pointed out in Section 2.

APPENDIX B

In this appendix, we find the most general solution

$$F_{mn}\equiv F(a_m,a_n,a_m^*,a_n^*) \tag{B.1}$$

satisfying equations (4.14), (4.16), and (4.19). Upon making the substitution

$$z_{mn}\equiv\ln(a_n^*F_{mn}) \tag{B.2}$$

equations (4.14) and (4.16) become, respectively,

$$z_{mn}=z_{nm}^* \tag{B.3}$$

$$\frac{1}{a_n}\frac{\partial}{\partial a_n^*}z_{mn}-\frac{1}{a_m^*}\frac{\partial}{\partial a_m}z_{mn}+\frac{1}{a_n^*}\frac{\partial}{\partial a_n}z_{mn}^*-\frac{1}{a_m}\frac{\partial}{\partial a_m^*}z_{mn}^*=0 \tag{B.4}$$

[We have used equation (B.3) to simplify equation (B.4).]

It is easily seen that if z_{mn} is an arbitrary function of $a_n, a_m^*, a_n^*a_n+a_m^*a_m$, then equation (B.4) is satisfied. If in addition to these three variables, we introduce a fourth variable $w\equiv a_n^*a_n-a_m^*a_m$, equation (B.4) becomes

$$\frac{\partial}{\partial w}(z_{mn}+z_{mn}^*)=0 \tag{B.5}$$

[3]The random phase assumption at $t=0$ does not imply that the phases are random at $t>0$ for the ordinary Schrödinger equation: equation (A.9a) fails in this case. The Schrödinger equation solutions seem to require a random phase approximation for all $t>0$, in order to be described by the Fokker–Planck equation (4.12).

Thus the most general solution of equation (B.4) can be written

$$z_{mn} = A(a_n, a_m^*, a_n^* a_n + a_m^* a_m) + iB \tag{B.6}$$

where A is an arbitrary complex function of its three arguments, and B is an arbitrary real function of all four variables.

The symmetry condition (B.3) requires

$$A = A_R(a_n a_m^*, a_n + a_m^*, a_n^* a_n + a_m^* a_m) \tag{B.7a}$$

where $A_R(\ ,\ ,\ )$ is a real-valued function when its arguments are real, and

$$B = \phi(s, u, v, w) - \phi(s, u, v, -w) \tag{B.7b}$$

where ϕ is a real function of its four real arguments $s \equiv a_n a_m^* + a_n^* a_m$, $u \equiv a_n a_m + a_n^* a_m^*$, $v \equiv a_n^* a_n + a_m^* a_m$, and w.

Thus, the most general expression for F_{mn} satisfying equations (4.14) and (4.16) is, from equations (B.2), (B.6), and (B.7):

$$F_{mn} = \frac{1}{a_n^*} R(a_n a_m^*, a_n + a_m^*, a_n^* a_n + a_m^* a_m) e^{i\{\phi(s,u,v,w) - \phi(s,u,v,-w)\}} \tag{B.8}$$

[we have written $R = \exp A_R$; $R(\ ,\ ,\ )$ is a real-valued function when its arguments are real].

For condition (4.19) to be obtained, $|R|^2 = 0$ must imply that $x_n x_m = 0$. Since

$$a_n a_m^* = 0 \Leftrightarrow x_n x_m = 0 \tag{B.9a}$$

$$a_n + a_m^* = 0 \not\Leftrightarrow x_n x_m = 0 \tag{B.9b}$$

$$a_n^* a_n + a_m^* a_m = 0 \Leftrightarrow x_n = 0 \quad \text{and} \quad x_m = 0 \tag{B.9c}$$

the only way to achieve (4.19) is for R to vanish when its first argument $a_n a_m^*$ vanishes [note that (B.9c) is too restrictive, since $x_n = 1$, $x_m = 0$ is a desired possibility]. The simplest way for this to happen is if $R \sim (a_n a_m^*)^r$, $r > 0$.

However, we must in fact have $r \geqslant 1$. This is because of the following technical point. As was mentioned in Section 2, in order to prove that the solution of a set of stochastic differential equations is a Markov process (and thereby satisfies a Fokker–Planck equation), the right-hand side of the stochastic equations must be bounded. Applying this to (4.4), we see that F_{mn} must be bounded. But by equation (B.8), F_{mn} will approach infinity as $|a_n^*| \to 0$ unless $R \sim a_n^r$, $r \geqslant 1$.

The simplest choice for F_{mn} is to cut out all dependence on variables other than the needed dependence on $a_n a_m^*$. We choose the lowest power of $a_n a_m^*$ consistent with all requirements, so

$$F_{mn} = \frac{a_n a_m^*}{a_n^*} \tag{B.10}$$

This has the important advantage of allowing us to write the reducing Schrödinger equation (2.4) as a quadratic form in the state vector $|\psi\rangle$.

REFERENCES

Ash, R., and Gardner, M. (1975). *Topics in Stochastic Processes*, Academic Press, New York.

Belinfante, F. (1973). *A Survey of Hidden-Variable Theories*, Pergamon, Oxford.

Belinfante, F. (1975). *Measurements and Time Reversal in Objective Quantum Theory*, Pergamon, Oxford.

Bialynicki-Birula, I., and Mycielski, J. (1976). *Ann. Phys.* (U.S.A.), **100**, 62.

Everett, H. (1957). *Rev. Mod. Phys.*, **24**, 454.

Friedman, A. (1975). *Stochastic Differential Equations and Applications*, Vols. I and II, Academic Press, New York.

Jauch, J. (1968). *Foundations of Quantum Mechanics*, Addison Wesley, Reading, Massachusetts.

Jauch, J. (1971). *Enrico Fermi School of Physics, 49th Course*, Academic Press, New York.

Kübler, O., and Zeh, H. (1973). *Ann. Phys. (N.Y.)*, **76**, 405.

McKean, H. P., Jr., (1969). *Stochastic Integrals*, Academic Press, New York.

McShane, E. J. (1960). *Proc. Natl. Acad. Sci. U.S.A.*, **63**, 275, 1084.

Papaliolios, C. (1967). *Phys. Rev. Lett.*, **18**, 622.

Pauli, W. (1926). *Festschrift zum 60 Geburtstag A. Sommerfelds*, Leipzig.

Pearle, P. (1967). *Am. J. Phys.*, **35**, 742.

Pearle, P. (1976). *Phys. Rev. D*, **13**, 857.

Schrödinger, E. (1935). *Naturwissenschaften*, **23**, 807.

Wang, M., and Uhlenbeck, G., (1945). *Rev. Mod. Phys.*, **17**, 323 (also in *Selected Papers on Noise and Stochastic Processes*, N. Wax, ed., Dover, New York, 1954).

Wong, E. (1971). *Stochastic Processes in Information and Dynamical Systems*, McGraw-Hill, New York.

Wong, E., and Zakai, M. (1965). *Int. J. Eng. Sci.*, **3**, 213.

Zeh, H. (1970). *Found. Phys.*, **1**, 69.

Zeh, H. (1971). *Enrico Fermi School of Physics, 49th Course*, Academic Press, New York.