PHYSICAL REVIEW A

GENERAL PHYSICS

THIRD SERIES, VOLUME 39, NUMBER 5 MARCH 1, 1989

Combining stochastic dynamical state-vector reduction with spontaneous localization

Philip Pearle

Department of Physics, Hamilton College, Clinton, New York 13323

(Received 6 October 1988)

A linear equation of motion for the state vector is presented, in which an anti-Hermitian Hamiltonian that fluctuates randomly is added to the usual Hamiltonian of the Schrödinger equation. It is shown how the resulting theory describes the continuous evolution of a state vector to an ensemble of reduced state vectors while retaining important physical features of the Ghirardi, Rimini, and Weber [Phys. Rev. D **34**, 470 (1986)] theory of spontaneous localization, in which the state-vector reduction occurs discontinuously. A novel aspect, compared with ordinary quantum theory, is that the state-vector norm changes with time. The squared norm of each state vector is interpreted as being proportional to the probability possessed by that state vector in the ensemble of state vectors. This interpretation is shown to be consistent with the independent Markovian evolution of each state vector.

I. DYNAMICAL REDUCTION

Dynamical reduction of the state vector is a process whereby a superposition such as

$$|\psi,0\rangle = a_1(0)|\psi_1\rangle + a_2(0)|\psi_2\rangle \tag{1}$$

continuously evolves with time into

$$|\psi,t\rangle = e^{i\theta_k}|\psi_k\rangle \quad (k=1 \text{ or } 2)\,. \tag{2}$$

A number of theories have so far been constructed[1-3] which modify the usual Schrödinger equation to achieve dynamical reduction. They enable one to interpret the state vector as corresponding to an individual situation in nature. A superposition of macroscopically distinguishable states cannot be interpreted in this way. The reduction is designed to prevent the existence, for any appreciable time interval, of such a superposition.

Stochastic dynamical reduction theories[2,3] add terms with randomly fluctuating coefficients to the Schrödinger equation, causing the squared amplitudes $X_k(t) \equiv |a_k(t)|^2$ to fluctuate randomly, although their sum $\sum_k X_k(t) = 1$ remains constant. Eventually one X_k reaches the value 1 and the other X_k's vanish.

The squared amplitudes X_k can be thought of as playing a (continuous in time) "gambler's ruin" game,[4] until one of the state vectors in the superposition "wins" the game. For example, the following game is analogous to the evolution (1)→(2). Two gamblers with initial "stakes" of $d_1(0), d_2(0)$ dollars toss a fair coin, and exchange a dollar depending on the outcome of each toss, until one of them wins all d_1+d_2 dollars. $d_k(t)/(d_1+d_2)$ is to be identified with $X_k(t)$, and a particular sequence of coin tosses is to be identified with a particular fluctuation of the coefficients.

It is easily shown,[5] because it is a *fair* game, that the average of $d_k(t)$ over the ensemble of games does not change with time and, as a result, the gambler with initial stake $d_k(0)$ has probability $d_k(0)/(d_1+d_2)$ of winning the game. It is precisely the analogous *Martingale* property

$$d\langle X_k(t)\rangle/dt = 0 \tag{3}$$

[the average $\langle\ \rangle$ is over the ensemble of trajectories of $X_k(t)$] that ensures agreement with the predictions of quantum theory: the outcome $X_k(t)\to 1$ in Eq. (2) occurs with probability $P_k = X_k(0)$. Here is the argument. From Eq. (3), $\langle X_k(0)\rangle = \langle X_k(t)\rangle$. On the one hand, $\langle X_k(0)\rangle = X_k(0)$ since we start with the pure state (1). On the other hand, for sufficiently large t so that Eq. (2) obtains for all state vectors in the ensemble $\langle X_k(t)\rangle = 0\times(\text{probability } X_k\to 0) + 1\times(\text{probability } X_k \to 1)$. Thus, $X_k(0) = P_k$.

Present dynamical reduction theories have concentrated on achieving the evolution (1)→(2), but have left important problems unanswered. Among these are the following problems.

(1) *The preferred basis problem.* What chooses the basis vectors which are the end products of reduction? Why does the reduction end in the macroscopic states we see around us [e.g., Eq. (2)], and not in a superposition of such states [e.g., $2^{-1/2}(|\psi_1\rangle \pm |\psi_2\rangle)$]?

(2) *The trigger problem.* What turns the reduction on? What mechanism is responsible for initiating a rapid reduction for a superposition of macroscopically distinguishable states while leaving a superposition of microscopically distinguishable states essentially unaffected?

II. SPONTANEOUS LOCALIZATION

Recently, Ghirardi, Rimini, and Weber[6] (GRW) (see also Bell[7]) have presented a "spontaneous localization" theory that deals with the preferred basis and trigger problems in an interesting way. Theirs is *not* a dynamical reduction theory: the state-vector alteration is presumed to occur instantaneously (we shall call this a "hit"), not as the result of a continuous evolution in time.

However, GRW give a prescription for what the state vector should be after a hit. If the wave function of a single particle before the hit is $\psi(\mathbf{x})$, after the hit it is multiplied by a Gaussian $\exp[-(\alpha/2)(\mathbf{x}-\mathbf{z})^2]$, and divided by a numerical factor N to correctly normalize the altered wave function. (GRW have chosen the localization width $\alpha^{-1/2}\approx 10^{-5}$ cm.) We shall call the point $\mathbf{z}$ the "center" of the hit. It is chosen randomly, but all centers are not equally likely: the frequency of hitting is presumed proportional to N^2. (GRW have selected the proportionality factor $\lambda\approx 10^{-16}$ s$^{-1}\approx 10^{-8}$ y^{-1}). Thus centers are most likely to appear where the wave function is largest.

A wave function well localized within $\alpha^{-1/2}$ will scarcely be affected by a hit. However, a wave function which is, for example, a superposition of two such localized packets separated by a distance $\gg\alpha^{-1/2}$ is most likely to be hit by a center near one packet, and that will dramatically reduce the size of the other packet in the resulting wave function. In this way (an excellent approximation to[8]) the reduction evolution (1)$\rightarrow$(2) is achieved. This resolves the preferred basis problem, for a single particle, in favor of spatially localized (within $\alpha^{-1/2}$) states.

The GRW response to the trigger problem is subtle. The trigger mechanism remains mysterious: no reason is given for the occurrence of a hit. It is presumed to be a universal process affecting all particles with equal likelihood. But, given this process, the different behavior of microscopic and macroscopic superpositions is cleverly explained. For a single particle, the hits occur so infrequently that no one has so far been able to think up an experiment precise enough to detect their presumed presence. But for a large object composed of many distinguishable particles, in a superposition of two spatially separated states (e.g., center of mass separation $\gg\alpha^{-1/2}$), a hit of a *single* particle will reduce the *whole* wave function. The frequency of such a hit is proportional to the number of particles in the body, so such a reduction will take place rapidly.

III. SYNTHESIS

In this article, we will present a stochastic dynamical reduction theory which resolves the preferred basis and trigger problems using the ideas of GRW. However, the Poisson process of instantaneous hits is replaced by a Markov process, with continuous evolution of the state vector. We shall call this a "continuous spontaneous localization" theory.

For simplicity, we will begin with an extensive discussion of a single particle moving in one dimension. This is sufficient to illustrate most of the novel features of this process. We propose the following equation of motion (an Itô stochastic differential equation) for the wave function:

$$d\psi(x,t)=-iH\psi(x,t)dt+[dw(x,t)-\tfrac{1}{2}\lambda\,dt]\psi(x,t)\ . \tag{4}$$

Equation (4) is remarkable, for a dynamical reduction theory, in that it is linear in ψ. H is the usual Hamilton of the Schrödinger equation. $w(x,t)$ is a real Brownian motion for each value of x. It is characterized by the following expectation values over the ensemble of Brownian motions:

$$\langle dw(x,t)\rangle=0\ , \tag{5a}$$

$$\langle dw(x,t)dw(x',t)\rangle=\lambda\Phi(x-x')dt\quad[\Phi(0)=1]\ . \tag{5b}$$

Thus the Brownian motions at different points of space are correlated.

The function Φ suggested by the GRW theory is

$$\Phi_{\mathrm{GRW}}(x-x')=\exp[-(\alpha/4)(x-x')^2] \tag{6}$$

but there are other suitable choices.[9] To make the connection with the GRW hits, and to facilitate later physical discussion, it is useful to write $dw(x,t)$ in terms of *uncorrelated* Brownian functions:

$$\langle dB(z,t)\rangle=0,\quad\langle dB(z,t)dB(z',t)\rangle=\delta(z-z')\lambda\,dt\ , \tag{7a}$$

$$dw(x,t)=\int dz\,dB(z,t)(\alpha/\pi)^{1/2}\exp[-(\alpha/2)(x-z)^2] \tag{7b}$$

[Eqs. (5) and (6) follow from Eqs. (7)].

In the remainder of this section we will give a qualitative discussion of the behavior of the solutions of Eq. (4): the quantitative justification for these remarks is in Secs. IV, and V, and VI. We will assume for simplicity that $H=0$, in order to concentrate on the effect of the new terms in Eq. (4).[10] Their most obvious feature is that they change the norm of the wave function.

Indeed, the last term in Eq. (4) acting by itself would cause the norm to exponentially decay with time constant λ. However, the randomly fluctuating term can increase or decrease the norm.

We will adopt the precept that the squared norm of each (unnormalized) wave function represents the weight associated with that (normalized) wave function in the ensemble of wave functions. This is a natural generalization of the GRW prescription that the frequency of a hit is proportional to the squared norm of the wave function after the hit.

It might appear that this precept makes the probability that a particular state vector is in the ensemble dependent upon the composition of the rest of the ensemble. This would be unacceptable, because an ensemble of independently evolving state vectors is an ensemble described by classical probability theory, in which there can be no in-

terference of probabilities. That this precept is, in fact, consistent with an independent Markovian evolution of each state vector is demonstrated in Appendix C. In what follows we shall take this consistency for granted. However, we must comment here upon how probabilities are to be calculated.

Usually when dealing with dynamics driven by Brownian motion there is, so to speak, one dynamical trajectory associated with each Brownian motion. That is not the case here. Let $d\Omega$ is the probability measure in the space of Brownian functions, and w_Ω be the particular Brownian function responsible for the evolution of the particular wave function ψ_Ω, with squared norm

$$N_\Omega^2(t) \equiv \int dx\, |\psi_\Omega(x,t)|^2 \ . \tag{8}$$

Then, according to our precept, the probability that ψ_Ω lies in the ensemble is $N_\Omega^2 d\Omega$, not $d\Omega$. Or course this interpretation requires, for consistency, that the sum of probabilities remains equal to 1,

$$\int N_\Omega^2(t) d\Omega \equiv \langle N^2(t) \rangle = 1 \ , \tag{9}$$

which we will see [Eq. (18a) and the following] follows from Eq. (4). We shall refer to the ensembles with weights $d\Omega$ and $N_\Omega^2 d\Omega$ as the *raw* and the *physical* ensembles, respectively. Now, what will be the behavior of the ensemble of norms, and which wave functions will predominate because their norms are largest?

Let us write out the essential part of Eq. (4) as

$$\psi(x, t+dt) = [1 + dw(x,t) - \tfrac{1}{2}\lambda\, dt]\psi(x,t) \tag{10a}$$

$$= \Big\{1 + \int dz\, dB(z,t)(\alpha/\pi)^{1/4} \times \exp[-(\alpha/2)(x-z)^2] - \tfrac{1}{2}\lambda\, dt\Big\}\psi(x,t) \ . \tag{10b}$$

One expects that, since dw (or dB) is as likely to fluctuate positively as negatively, the randomly fluctuating term will have only a modest effect for the majority of Brownian motions in the raw ensemble. Therefore the norms of the wave functions evolving subject to these Brownian motions will decrease roughly exponentially due to the last term in the bracket of Eq. (10). Thus these wave functions will be essentially unaffected in shape, but after a few time constants they will be of negligible norm and of neglible importance in the physical ensemble.

On the other hand, for that minority of Brownian motions for which $dw(x,t)$, or a set of $dB(z,t)$ with $x - \alpha^{-1/2} \le z \le x + \alpha^{-1/2}$, happens to be positive significantly more often than it is negative, the norm of the associated wave functions will grow *if x also* happens to have an $\alpha^{-1/2}$ neighborhood where $\psi(x)$ is large. This can be seen most easily from Eq. (10b), where the explicit appearance of the Gaussian shows that, e.g., a sequence of positive increments of $dB(z,t)$ will increase the amplitude of the wave function in the neighborhood of $x = z$, at the expense of the amplitude of the wave function for x outside the $\alpha^{-1/2}$ neighborhood of z.

Of course, this can also be seen from (the equivalent) Eq. (10a) if we remember that a sequence of positive increments of dw at x implies also a sequence of positive increments for locations within the $\alpha^{-1/2}$ neighborhood of x, because of the correlation (5b). Conversely, the increments in dw outside the $\alpha^{-1/2}$ neighborhood of x are uncorrelated with those inside, and are most likely to be equally positive and negative. Therefore, if ψ should grow in the $\alpha^{-1/2}$ neighborhood of x, it is most likely that there will be a concomitant approximate exponential decrease of ψ outside that neighborhood.

After a few time constants, the ensemble that results from a wave function which initially is a superposition of two packets each of width $\ll \alpha^{-1/2}$ separated by a distance $\gg \alpha^{-1/2}$ (we shall call this the "canonical" wave function hereafter), is as follows. In the raw ensemble, the huge majority of wave functions still contain two packets, but their norms have decreased exponentially. There is also a small set of wave functions which consist essentially of just one packet, with very large norms. On the other hand, in the physical ensemble the weighting is reversed. The measure of wave functions containing two packets is negligibly small. The overwhelmingly probable wave functions in the ensemble consist of essentially just one packet.

Now, what does the theory predict for the probability associated with the wave functions containing a single packet? The prediction is identical to that of quantum theory: in the physical ensemble, the description of the evolution of the two packets turns out to be the gambler's ruin description. In particular, using the notation of Sec. I [e.g., $|\psi_1\rangle$ and $|\psi_2\rangle$ in Eqs. (1) and (2) represent the two packets], the squared amplitudes $X_k(t)$ possess the Martingale property (3). From this follows agreement with the predictions of quantum theory, as we have seen in Sec. I.

We may understand the origin of this crucial agreement by anticipating a result [Eq. (18a)] of Sec. IV. As a consequence of Eq. (4) (with $H = 0$), the individual squared amplitudes at each point x are a Martingale:

$$d\langle |\psi(x,t)|^2 \rangle / dt = 0 \ . \tag{11}$$

To see how Eq. (11) is responsible for producing agreement with the predictions of quantum theory, define the norms $N_{\Omega 1}$ and $N_{\Omega 2}$ of each separate packet (occupying regions 1 or 2) belonging to an individual unnormalized wavefunction $\psi_\Omega(x,t)$,

$$N_{\Omega k}^2(t) \equiv \int_k dx\, |\psi_\Omega(x,t)|^2 \ . \tag{12}$$

By integrating Eq. (11) over region 1 or 2, and using the definition (12), we see that

$$d\langle N_k^2(t) \rangle / dt = 0 \ . \tag{13}$$

Now, the squared norm associated with ψ_Ω is $N_\Omega^2(t) = N_{\Omega 1}^2(t) + N_{\Omega 2}^2(t)$ [Eq. (8)]. However, for $t = T$ greater than a few time constants, either $N_{\Omega 1}^2$ or $N_{\Omega 2}^2$ or both essentially vanish as we have discussed. Therefore the probability P_k, in the physical ensemble, that a wave function consists of just the kth packet, for $t \ge T$, is

$$P_k = \int_k N_{\Omega k}^2(T) d\Omega = \langle N_k^2(T) \rangle \ . \tag{14}$$

It follows from Eq. (13) that $\langle N_k^2(0)\rangle = \langle N_k^2(T)\rangle$. But since $\langle N_k^2(0)\rangle = X_k(0)$, we obtain from Eq. (14) the quantum theory prediction $P_k = X_k(0)$.

IV. SINGLE PARTICLE: BEHAVIOR AT A POINT

We now wish to justify the statements of the previous section by analyzing the consequences of Eq. (4) (with $H=0$). The phase of $\psi(x,t)$ is unaffected by the evolution (4) because w is real.[11] To see how the unnormalized squared amplitudes $Y(x,t)\equiv|\psi(x,t)|^2$ behave in the raw ensemble (all we consider in this section) we utilize the Itô formula

$$d(fg)=f\,dg+g\,df+\langle df\,dg\rangle\ , \tag{15}$$

where $f=\psi(x,t)$ and $g=\psi^*(x',t)$, to obtain from Eq. (4)

$$\begin{aligned} d[\psi(x,t)\psi^*(x',t)]=\{&-i[H(x)-H(x')]dt \\ &+[dw(x,t)+dw(x',t)] \\ &+\lambda[\Phi(x-x')-1]dt\} \\ &\times\psi(x,t)\psi^*(x',t) \end{aligned} \tag{16}$$

(which will be useful later, when we discuss the density matrix in Sec. VII) and, setting $x=x'$ and $H=0$,

$$dY(x,t)=2\,dw(x,t)Y(x,t)\ . \tag{17}$$

Equations (17) and (5) contain the complete information about the ensemble of wave functions.

First, consider the behavior of $Y(x,t)$ at a single point x, according to the raw ensemble description. From Eq. (17) we can calculate the moments

$$\langle dY(x,t)\rangle=0\ , \tag{18a}$$

$$\langle[dY(x,t)]^2\rangle=4\lambda Y(x,t)^2dt\ . \tag{18b}$$

[Note that Eq. (11) follows from Eq. (18a) while Eq. (9) follows from integrating Eq. (11) over x.] Using Eqs. (18a) and (18b) we can construct the Fokker-Planck equation describing the behavior of the ensemble of squared amplitudes $Y(x)$ by means of the prescription

$$\begin{aligned} \frac{\partial\rho(Y;t)}{\partial t}=&-\left(\frac{\partial}{\partial Y}\right)\left[\left(\frac{\langle dY\rangle}{dt}\right)\rho\right] \\ &+\tfrac{1}{2}\left(\frac{\partial}{\partial Y}\right)^2\left[\left(\frac{\langle(dY)^2\rangle}{dt}\right)\rho\right]\ . \end{aligned} \tag{19}$$

$\rho(Y;t)dY$ is the probability that Y lies in the range dY in the raw ensemble.

Putting Eqs. (18) into Eq. (19) we obtain

$$\partial\rho(Y;t)/\partial t=2\lambda(\partial/\partial Y)^2[Y^2\rho(Y;t)]\ . \tag{20}$$

The solution of Eq. (20), subject to the initial condition $\rho(Y;0)=\delta(Y-Y_0)$ is

$$\begin{aligned} &\rho(Y;t)dY \\ &=(8\pi\lambda t)^{-1/2}\{\exp[-(\ln Y-\ln Y_0+2\lambda t)^2/8\lambda t]\}d\ln Y\ . \end{aligned} \tag{21}$$

Equation (21) shows that $\rho(0;t)=0$, and $\rho(Y;t)$ rises to a peak at $Y=Y_0\exp(-2\lambda t)$, falling to zero as $Y\to\infty$. Thus, in the raw ensemble, most wave functions have squared amplitudes which decrease roughly exponentially with time at each point. However, there is a tail in the distribution, so some wave functions acquire very large squared amplitudes.

Next consider the correlated behavior of $Y(x_1,t), Y(x_2,t)$ at two different points x_1,x_2. From Eq. (17) we find

$$\langle dY(x_1,t)Y(x_2,t)\rangle=4\lambda\Phi(x_1-x_2)Y(x_1,t)Y(x_2,t)dt\ . \tag{22}$$

Putting (22) and (18) into the generalization of Eq. (19) to more than one variable

$$\begin{aligned} \partial\rho(\mathbf{Y};t)/\partial t=&-\sum_n(\partial/\partial Y_n)[(\langle dY_n\rangle/dt)\rho] \\ &+\tfrac{1}{2}\sum_{n,m}(\partial/\partial Y_n)(\partial/\partial Y_m) \\ &\times[(\langle dY_n dY_m\rangle/dt)\rho]\ , \end{aligned} \tag{23}$$

we obtain, for an arbitrary number of points,

$$\begin{aligned} \partial\rho(\mathbf{Y};t)/\partial t=&2\lambda\sum_{n,m}\Phi(x_n-x_m)(\partial/\partial Y_n)(\partial/\partial Y_m) \\ &\times Y_nY_m\rho\ , \end{aligned} \tag{24a}$$

which, for two points, is

$$\begin{aligned} \partial\rho(\mathbf{Y};t)/\partial t=&2\lambda[(\partial/\partial Y_1)^2Y_1^2+(\partial/\partial Y_2)^2Y_2^2 \\ &+2\Phi(x_1-x_2)(\partial/\partial Y_1)(\partial/\partial Y_2) \\ &\times Y_1Y_2]\rho\ . \end{aligned} \tag{24b}$$

We see from Eq. (24b), when x_1 is distant from x_2 (i.e., $|x_1-x_2|\gg\alpha^{-1/2}$, so $\Phi\approx0$), the evolutions of Y_1 and Y_2 are uncorrelated, and the raw probability distribution is the product of separate distributions (21) for Y_1, Y_2. On the other hand, for nearby points ($|x_1-x_2|\ll\alpha^{-1/2}$, so $\Phi\approx1$), the solution of Eq. (24b) subject to the initial condition $\rho(Y_1,Y_2;0)=\delta(Y_1-Y_{10})\delta(Y_2-Y_{20})$ is

$$\begin{aligned} \rho\,dY_1dY_2=&\delta[\ln(Y_1/Y_2)-\ln(Y_{10}/Y_{20})](32\pi\lambda t)^{-1/2} \\ &\times\exp\{[-(32\lambda t)^{-1}] \\ &\times[\ln Y_1Y_2-\ln Y_{10}Y_{20}+4\lambda t]^2\} \\ &\times d\ln Y_1\,d\ln Y_2\ . \end{aligned} \tag{25}$$

Because of the δ function in Eq. (25), the *ratio* of squared amplitudes at nearby points remains constant. This conclusion is also reached using the physical ensemble.

This feature, that the shape of each wave function in a $\alpha^{-1/2}$ neighborhood of each point x remains relatively undisturbed, is worth emphasizing. To examine this behavior more precisely, without making the approximation $\Phi\approx1$ employed above, we write down the equation of motion for the ratio Y_1/Y_2. Using Eq. (15) with $f=Y_1$ and $g=1/Y_2$, together with the Itô formula

$$dg(Y)=g'dY+\tfrac{1}{2}g''\langle(dY)^2\rangle \tag{26}$$

and Eq. (17), we obtain

$$d(Y_1/Y_2)=(2[dw(x_1,t)-dw(x_2,t)] +4\lambda\{1-\exp[-(\alpha/4)(x_1-x_2)^2]\}dt) \times(Y_1/Y_2)\ . \quad (27)$$

From Eq. (27) we pluck out the drift and diffusion of Y_1/Y_2,

$$\langle d(Y_1/Y_2)\rangle/dt =4\lambda\{1-\exp[-(\alpha/4)(x_1-x_2)^2]\}(Y_1/Y_2)\ , \quad (28)$$

$$\langle [d(Y_1/Y_2)]^2\rangle/dt =4\langle [dw(x_1,t)-dw(x_2,t)]^2\rangle/dt =8\lambda\{1-\exp[-(\alpha/4)(x_1-x_2)^2]\}(Y_1/Y_2)^2\ . \quad (29)$$

According to Eqs. (28) and (29), both drift and diffusion of Y_1/Y_2 are proportional to λ for $|x_1-x_2|\gg\alpha^{-1/2}$, while for $|x_1-x_2|\ll\alpha^{-1/2}$ they are proportional $\lambda\alpha(x_1-x_2)^2\ll\lambda$. In this way the global wave function is altered dramatically on a time scale of order λ^{-1} while locally it is scarcely affected on the same time scale.

V. ONE PARTICLE: BEHAVIOR OF PACKETS

Now we turn to the canonical situation of an initial wave function consisting of two well-separated packets. As in Eq. (12), we define the norm $N_k(t)$ of each packet. From (17) we find how the norms change with time,

$$dN_k(t)^2=d\int_k dx\ Y(x,t)=2\int_k dw(x,t)Y(x,t)dx\ , \quad (30)$$

by means of which we can calculate the drift and diffusion of N_1^2,N_2^2,

$$\langle dN_1^2\rangle/dt=0,\quad \langle dN_2^2\rangle/dt=0\ , \quad (31a)$$

$$\langle dN_k^2 dN_{k'}^2\rangle=4\lambda\int_k dx\int_{k'}dx'\Phi(x-x')Y(x,t)Y(x',t)dt\ . \quad (31b)$$

For $k=k'$ we have $\Phi(x-x')\approx 1$, while for $k\neq k'$ we have $\Phi(x-x')\approx 0$, so from (31b)

$$\langle (dN_k^2)^2\rangle\approx 4\lambda(N_k^2)^2dt,\quad \langle dN_1^2 dN_2^2\rangle\approx 0\ . \quad (31c)$$

Thus the squared norms of the two packets are uncorrelated. In fact, by Eqs. (31a) and (31c) the Fokker-Planck equation and its solution describing the distribution of N_k^2 in the raw ensemble is given by Eqs. (20) and (21) (N_k^2 replaces Y): the squared norms of the two packets behave precisely the same way as did the squared amplitudes at two distant points.

We now are in a position to discuss the evolution of the joint probability distribution for N_1^2,N_2^2. In the *raw* ensemble, the probability density $\rho(N_1^2,N_2^2;t)$ is given by

$$\rho(N_1^2,N_2^2;t)dN_1^2dN_2^2=\rho(N_1^2;t)\rho(N_2^2;t)dN_1^2dN_2^2=(8\pi\lambda t)^{-1}d\ln N_1^2 d\ln N_2^2 \times\exp[-(8\lambda t)^{-1}]\{[\ln(N_1^2/N_{10}^2)+2\lambda t]^2+[\ln(N_2^2/N_{20}^2)+2\lambda t]^2\} \quad (32a)$$

[using Eq. (21) for $\rho(N_k^2;t)$]. In the *physical* ensemble, the probability density is $r(N_1^2,N_2^2;t)\equiv(N_1^2+N_2^2)\rho(N_1^2,N_2^2;t)$ which may be written as

$$rdN_1^2dN_2^2=(N_{10}^2\exp[-(8\lambda t)^{-1}]\{[\ln(N_1^2/N_{10}^2)-2\lambda t]^2+[\ln(N_2^2/N_{20}^2)+2\lambda t]^2\} +N_{20}^2\exp[-(8\lambda t)^{-1}]\{[\ln(N_1^2/N_{10}^2)+2\lambda t]^2+[\ln(N_2^2/N_{20}^2)-2\lambda t]^2\})(8\pi\lambda t)^{-1}d\ln N_1^2 d\ln N_2^2\ . \quad (32b)$$

Equations (32a) and (32b) display the dramatic difference between the physical and raw ensembles.

The raw ensemble probability density (32a) is the product of two probability densities, in each of which there is a peak which moves toward $N_k^2=0$. The physical ensemble probability density (32b) is the sum of two terms. Each term is itself the product of two probability densities, in one of which the peak moves toward $N_1^2=\infty$ ($N_2^2=\infty$) while the other peaks moves toward $N_2^2=0$ ($N_1^2=0$). Moreover, the overall probability associated with each term is $N_{10}^2=X_1(0)$ [$N_{20}^2=X_2(0)$], the squared amplitude for the packet in the original wave function. Thus Eq. (32b) shows explicitly the reduction behavior, how the initial wave function with probability density $r(N_1^2,N_2^2;0)=\delta(N_1^2-N_{10}^2)\delta(N_2^2-N_{20}^2)$ evolves into an ensemble, with correct probability distribution, of wave functions of ever-increasing norm, each containing (essentially) just one packet.

VI. GAMBLER'S RUIN

In the previous sections we discussed the *unnormalized* wave functions. We described the behavior of the squared amplitudes $Y(x,t)$ and the behavior of the norms of localized packets $N_k(t)$. In this section we comment on the behavior of the squared amplitudes and packet norms of *normalized* wave functions, using the correct probability densities of the physical ensemble.

The Fokker-Planck equation for the probability density of the squared amplitudes

$$Z(x,t)\equiv|\psi(x,t)|^2/\int dx'|\psi(x',t)|^2 \quad (33)$$

is obtained in Appendix A, Eq. (A11). That the $Z(x,t)$ for different x play a gambler's ruin game among themselves is discussed in Appendix A, and we will say no more about this here.

Now let us consider the canonical two-packet situa-

tion. We shall use the result (32b) of Sec. V to find the probability density distribution for the squared norm of packet 1,

$$X_1(t)\equiv N_1^2(t)/[N_1^2(t)+N_2^2(t)] \tag{34}$$

(X_2's distribution is implied since $X_2=1-X_1$). This probability density $R(X_1;t)$ is obtained by integrating $r(N_1^2,N_2^2;t)$ over all norms $N^2\equiv N_1^2+N_2^2$,

$$R(X_1;t)=\int_0^\infty dN^2 r(X_1N^2,(1-X_1)N^2;t) \times J(N_1^2,N_2^2|X_1,N^2)\ . \tag{35a}$$

J is the Jacobian determinant, whose value is calculated from Eq. (34) to be N^2. Upon substituting the expression (32b) for r into Eq. (34a), and performing the integral, we obtain

$$R(X_1;t)dX_1=\{X_{10}\exp[-(16\lambda t)^{-1}][-4\lambda t+\ln X_1/(1-X_1)-\ln X_{10}/(1-X_{10})]^2 +(1-X_{10})\exp[-(16\lambda t)^{-1}][4\lambda t+\ln X_1/(1-X_1)-\ln X_{10}/(1-X_{10})]^2\}(16\pi\lambda t)^{-1}d\ln X_1/(1-X_1)\ . \tag{35b}$$

Equation (35b) describes a solution whose initial distribution $\delta(X_1-X_{10})$ breaks up on a time scale λ^{-1} into two peaks which travel toward $X_1=1$ and $X_1=0$, the areas under the peaks being x_{10} and $1-X_{10}$, respectively. It may be directly verified that Eq. (35b) is the solution of the Fokker-Planck equation

$$\partial R(X_1;t)/\partial t=4\lambda(\partial/\partial X_1)^2[X_1^2(1-X_1)^2R(X_1;t)]\ . \tag{36}$$

In Appendix A, Eq. (A17), we give the generalization of Eq. (36) to any number of packets

$$\partial R(\mathbf{X};t)/\partial t=2\lambda\sum_{j,k}\left[\frac{\partial}{\partial X_j}-\frac{\partial}{\partial X_k}\right]^2 X_jX_k \times\left[X_j+X_k-\sum_i X_i^2\right]R(\mathbf{X};t)\ . \tag{37}$$

The Martingale property Eq. (3), the constancy of $\sum_i X_i$, and therefore the gambler's ruin nature of the competition between the X_k's follows immediately from Eq. (37).

Equations (35b), (36), and (37) are not new in the history of stochastic dynamical reduction theories. Some time ago we proposed[2] a theory whose two-state sector we showed is described by Eqs. (35b) and (36). More recently, Gisin[3] has suggested another theory whose two-state sector is also described by Eqs. (35b) and (36). We have proved[12] that Eq. (37) is the unique description of stochastic dynamical reduction theories with nonevolving phase angles whose off-diagonal density matrix elements decay exponentially with a universal time constant, that this behavior is a necessary condition for there to be no superluminal communication via the reduction mechanism, and that Gisin's theory is described by Eq. (37). We have also emphasized that a superposition described by Eq. (37) is never completely reduced.[8]

VII. DENSITY MATRIX

The probability density description of the behavior of the ensemble of wave functions is a complete description. In particular, it enables one to calculate the quantum-mechanical density matrix which, although it contains much less information than the probability density, is a sufficient tool for making all experimental predictions.

In some dynamical reduction theories the equation of time evolution of the density matrix does not depend upon the density matrix alone. One is then obliged to use the probability density to construct the density matrix and thus to make predictions. However, this is not the case here, as will be shown.

To construct the density matrix from an individual unnormalized wave function $\psi_\Omega(x,t)$, it is necessary to normalize it first, obtaining $\psi_\Omega(x,t)/N_\Omega(t)$. The probability that this normalized wave function is in the physical ensemble is $d\Omega\, N_\Omega^2(t)$. Therefore the density matrix is

$$D(x,x';t)\equiv\int d\Omega\, N_\Omega^2(t)[\psi_\Omega(x,t)/N_\Omega(t)] \times[\psi_\Omega^*(x',t)/N_\Omega(t)] =\langle\psi(x,t)\psi^*(x',t)\rangle\ . \tag{38}$$

Thus the density matrix can be found from the expectation value of $\psi(x,t)\psi^*(x',t)$ calculated in the raw ensemble. As a direct consequence of the equation of motion (4) for $d\psi$ we have already found the expression (16) for $d[\psi(x,t)\psi^*(x',t)]$. Taking the expectation value of Eq. (16) we obtain

$$\partial D(x,x';t)/\partial t=\{-i[H(x)-H(x')] +\lambda[\Phi(x-x')-1]\}D(x,x';t)\ . \tag{39}$$

Equation (39) clearly shows how the usual Hamiltonian evolution occurs for density matrix elements in the position representation taken between nearby points ($|x-x'|\ll\alpha^{-1/2}$, $\Phi\approx 1$), but is modified by the exponential decay of off-diagonal matrix elements between distance points ($|x-x'|\gg\alpha^{-1/2}$, $\Phi\approx 0$). Of course, Eq. (39) can also be obtained from the probability density [see Appendix A, Eqs. (A12) and (A13)].

Equation (39) is identical to the equation of evolution of the density matrix proposed by GRW.[6] However, we wish to emphasize that the behavior of the ensemble of wave functions subject to the GRW Poisson process is different from their behavior subject to the Markov process described here, even though their density matrices are identical at every instant of time.

To illustrate, consider the canonical two-packet situation. The Markov process probability density $R(X_1,X_{10};t)$ is given by Eq. (35b). The Poisson process probability density $R_{\mathrm{GRW}}(X_1,X_{10};t)$ is obtained in Appendix B, Eq. (B13), as

$$R_{\mathrm{GRW}}(X_1,X_{10};t)\approx\delta(X_1-X_{10})\exp(-\lambda t)+R(X_1,X_{10};\alpha(x_2-x_1)^2)[1-\exp(-\lambda t)] \tag{40a}$$

$$\approx\delta(X_1-X_{10})\exp(-\lambda t)+[X_{10}\delta(1-X_1)+(1-X_{10})\delta(X_1)][1-\exp(-\lambda t)]\ . \tag{40b}$$

Both probability densities R and R_{GRW} have initial value $\delta(X_1-X_{10})$ and final value $\approx X_{10}\delta(1-X_1)+(1-X_{10})\delta(X_1)$. However, for $t\ll\lambda^{-1}$, R describes an ensemble in which a negligible amount of wave-function reduction has occurred, i.e., there is a negligible probability that $X_1\approx 1$ or 0. It takes a few time constants for the initial probability density peak to diffuse toward $X_1=1$ or 0 and build up significant probability concentrations at these locations. On the other hand, R_{GRW} describes the immediate creation of reduced wave functions [last term in Eq. (40b) $\sim\lambda t$] for small t.

VIII. DISTINGUISHABLE PARTICLES

The generalization of Eqs. (4) and (5) to n distinguishable particles moving in three-dimensional space is

$$d\psi(\mathbf{x}_1,\ldots,\mathbf{x}_n;t)=-iH\psi\,dt+\left[\sum_k dw_k(\mathbf{x}_k,t)-(\tfrac{1}{2}n\lambda\,dt)\right]\psi\ , \tag{41}$$

$$\langle dw_k(\mathbf{x}_k,t)\rangle=0\ , \tag{42a}$$

$$\langle dw_k(\mathbf{x}_k,t)dw_l(\mathbf{x}'_l,t)\rangle=\delta_{kl}\lambda\Phi(|\mathbf{x}_k-\mathbf{x}'_l|)dt\ . \tag{42b}$$

First we look at the density matrix evolution equation. $d\psi\psi^*$ can be calculated from Eq. (41) just as Eq. (16) was obtained from Eq. (4) in Sec. IV,

$$d[\psi(\mathbf{x}_1,\ldots,\mathbf{x}_n;t)\psi^*(\mathbf{x}'_1,\ldots,x'_n;t)]=\left[-i[H(x)-H(x')]dt+\sum_k[dw_k(\mathbf{x}_k,t)+dw_k(\mathbf{x}'_k,t)]\right.$$
$$\left.+\lambda\left[\sum_k\Phi(|\mathbf{x}_k-\mathbf{x}'_k|)-n\right]dt\right]\psi(x,t)\psi^*(x',t)\ . \tag{43}$$

Remembering the result of Sec. VII,

$$D(\mathbf{x}_1,\ldots,\mathbf{x}_n,\mathbf{x}'_s,\ldots,x'_n;t)$$
$$=\langle\psi(\mathbf{x}_1,\ldots,\mathbf{x}_n;t)\psi^*(\mathbf{x}'_1,\ldots,\mathbf{x}'_n;t)\rangle\ ,$$

by taking the expectation value of Eq. (43) we obtain the density matrix evolution equation

$$\partial D(\mathbf{x},\mathbf{x}';t)/\partial t=\left[-i[H(\mathbf{x})-H(\mathbf{x}')]\right.$$
$$\left.+\lambda\left[\sum_k\Phi(|\mathbf{x}_k-\mathbf{x}'_k|)-n\right]\right]$$
$$\times D(\mathbf{x},\mathbf{x}';t)\ . \tag{44}$$

Equation (44) is precisely the density matrix evolution equation proposed by GRW.[6] Following them, consider an initial wave function describing a macroscopic body whose center of mass $\mathbf{R}\equiv n^{-1}\sum_k\mathbf{x}_k$ (we take all masses equal for simplicity) has an arbitrary probability distribution, but whose relative coordinates $\mathbf{r}_k\equiv\mathbf{x}_k-\mathbf{R}$ are each well localized within a spherical volume of radius $\ll\alpha^{-1/2}$. Since $\Phi(|\mathbf{x}_k-\mathbf{x}'_k|)=\Phi(|\mathbf{r}_k-\mathbf{r}'_k+\mathbf{R}-\mathbf{R}'|)$, and when $|\mathbf{R}-\mathbf{R}'|$ is comparable in size to $|\mathbf{r}_k-\mathbf{r}'_k|$ we have $\Phi\approx 1$, we may replace each $\Phi(|\mathbf{x}_k-\mathbf{x}'_k|)$ in Eq. (44) with $\Phi(|\mathbf{R}-\mathbf{R}'|)$ to a good approximation. Then Eq. (44) (with $H=0$) becomes

$$\partial D(\mathbf{r},\mathbf{R},\mathbf{r}',\mathbf{R}';t)/\partial t$$
$$\approx n\lambda[\Phi(|\mathbf{R}-\mathbf{R}'|)-1]D(\mathbf{r},\mathbf{R},\mathbf{r}',\mathbf{R}';t)\ . \tag{45}$$

Equation (45) expresses GRW's result that for a macroscopic body there is a rapid reduction [time constant $(\lambda n)^{-1}$] in the center of mass, while there is essentially no effect on the relative motion. This can be seen by taking the trace of Eq. (45) over the relative coordinates $\mathbf{r}_k$ or over the center of mass coordinate $\mathbf{R}$: the reduced density matrix $D(\mathbf{R},\mathbf{R}';t)$ satisfies Eq. (45), while the other reduced density matrix satisfies $\partial D(\mathbf{r},\mathbf{r}';t)/\partial t=0$.

Of course these results for the density matrix have their counterparts in the probability density description of the ensemble of wave functions. Let us examine the behavior of the squared amplitudes of the unrenormalized wave function $Y(\mathbf{x}_1,\ldots,\mathbf{x}_n;t)\equiv|\psi(\mathbf{x}_1,\ldots,\mathbf{x}_n;t)|^2$ just as was done for a single particle moving in one dimension in Sec. IV. Setting $\mathbf{x}_k=\mathbf{x}'_k$ (and $H=0$) in Eq. (43) results in the basic equation of motion for Y,

$$dY(\mathbf{x}_1,\ldots,\mathbf{x}_n;t)=2\sum_k dw_k(\mathbf{x}_k,t)Y(\mathbf{x}_1,\ldots,\mathbf{x}_n;t)\ , \tag{46}$$

while the approximations made prior to Eq. (45) convert Eq. (42b) to

$$\langle dw_k(\mathbf{x}_k,t)dw_l(\mathbf{x}'_l,t)\rangle\approx\delta_{kl}\lambda\Phi(|\mathbf{R}-\mathbf{R}'|)dt\ . \tag{47}$$

Therefore, from Eqs. (46) and (47), the diffusion coefficient for $Y(\mathbf{r}_1,\ldots,\mathbf{r}_{n-1},\mathbf{R};t)\equiv Y(\mathbf{x}_1,\ldots,\mathbf{x}_n;t)$ is

$$\langle dY(\mathbf{r},\mathbf{R};t)dY(\mathbf{r}',\mathbf{R}';t)\rangle/dt$$
$$\approx 4\lambda n\Phi(|\mathbf{R}-\mathbf{R}'|)Y(\mathbf{r},\mathbf{R};t)Y(\mathbf{r}',\mathbf{R}';t)\ , \tag{48}$$

which is all one needs (to construct the Fokker-Planck equation and therefore) to describe the complete behavior of the ensemble of squared amplitudes Y.

First let us look at the behavior of the center of mass.

By integrating Eq. (48) over the relative coordinates $\mathbf{r}_k, \mathbf{r}'_k$, $k=1,\ldots,n-1$ we obtain

$$\langle dY_{\text{c.m.}}(\mathbf{R};t)Y_{\text{c.m.}}(\mathbf{R}';t)\rangle \approx 4\lambda n\Phi(|\mathbf{R}-\mathbf{R}'|)Y_{\text{c.m.}}(\mathbf{R};t)Y_{\text{c.m.}}(\mathbf{R}';t)dt\ , \quad (49)$$

where $Y_{\text{c.m.}}(\mathbf{R};t)=\int d\mathbf{r}_1\ldots d\mathbf{r}_{n-1}Y(\mathbf{r},\mathbf{R};t)$ is the squared amplitude for the center of mass alone.

Equation (49) is identical in form to the diffusion coefficient for a single particle given in Eq. (22), except that λ in Eq. (22) is replaced by λn, and one spatial dimension is replaced by three. Therefore the whole single particle behavior discussed in Secs. IV–VI and Appendix A holds for the center of mass, except that reduction of the center of mass in the canonical two-packet state proceeds more rapidly, with time constant $(\lambda n)^{-1}$.

To see that the relative coordinates are not disturbed, one need only employ Eq. (46) to calculate the diffusion and drift of $Y(\mathbf{r},\mathbf{R};t)/Y(\mathbf{r}',\mathbf{R}';t)$ as was done in Sec. IV. The result has the same form as Eqs. (28) and (29),

$$\langle d(Y/Y')\rangle = 4\lambda n[1-\Phi(|\mathbf{R}-\mathbf{R}'|)]dt\ , \quad (50a)$$

$$\langle [d(Y/Y')]^2\rangle = 8\lambda n[1-\Phi(|\mathbf{R}-\mathbf{R}'|)]dt\ . \quad (50b)$$

For $\mathbf{R}=\mathbf{R}'$ we see that both drift and diffusion coefficients (50) vanish, so there is no change in the shape of the wave function in the relative coordinates.

Extending the GRW Poisson process to many indistinguishable particles has proved to be not an easy task.[13] The utility of the Markov process described here makes itself apparent in the ease with which this extension is accomplished within its framework. The generalization of Eq. (4) or Eq. (41) to many indistinguishable particles is

$$d\langle \mathbf{x}_1,\ldots,\mathbf{x}_n|\psi,t\rangle = \left[-iH+\sum_k dw(\mathbf{x}_k,t)-\tfrac{1}{2}\left\langle\left[\sum_k dw(\mathbf{x}_k,t)\right]^2\right\rangle\right]\langle \mathbf{x}_1,\ldots,\mathbf{x}_n|\psi,t\rangle \quad (51a)$$

$$= \left[-iH+\sum_k dw(\mathbf{x}_k,t)-\tfrac{1}{2}\lambda\, dt\sum_{kl}\Phi(|\mathbf{x}_k-\mathbf{x}_l|)\right]\langle \mathbf{x}_1,\ldots,\mathbf{x}_n|\psi,t\rangle\ . \quad (51b)$$

The statistical properties of $dw(\mathbf{x},t)$ are described by Eq. (5) as usual [with $(x-x')$ replaced by $|\mathbf{x}-\mathbf{x}'|$]. Unlike Eq. (41) for distinguishable particles, only one Brownian function $w(\mathbf{x}_k,t)$ is needed.

The calculation of the evolution equation for the density matrix proceeds as in Secs. VII and VIII. One calculates $d[\psi(\mathbf{x};t)\psi^*(\mathbf{x}';t)]$ using Eq. (51) and the Itô rule (15) and, upon taking the expectation value, one obtains

$$\partial D(\mathbf{x}_1,\ldots,\mathbf{x}_n,\mathbf{x}'_1,\ldots,\mathbf{x}'_n;t)/\partial t = \Big[-i[H(\mathbf{x})-H(\mathbf{x}')]+\lambda\sum_{kl}[\Phi(|\mathbf{x}_k-\mathbf{x}'_l|)-\tfrac{1}{2}\Phi(|\mathbf{x}_k-\mathbf{x}_l|) -\tfrac{1}{2}\Phi(|\mathbf{x}'_k-\mathbf{x}'_l|)]\Big]D(\mathbf{x},\mathbf{x}';t)\ . \quad (52)$$

After having tantalizingly displayed the basic equations (51) and (52) of the theory for many indistinguishable particles, we will refrain from drawing the interesting physical consequences here. They will be discussed in a forthcoming paper.[14]

IX. CONCLUDING REMARKS

The equations of motion presented here [Eqs. (4) or (41) or (51)] describe a nonunitary but linear evolution of the state vector. Subject to a particular Brownian function $w_\Omega(\mathbf{x},t)$, a sum of two state vectors at time 0 becomes, at time t, a state vector which is the sum of the two evolved state vectors. However, the theory is nonlinear in its rule of weighting each state vector's importance in the ensemble by the squared norm of the state vector. In these respects it is similar to ordinary quantum theory, with its linear evolution equation and nonlinear probabilistic interpretation.

The linearity of the equation of motion should prove useful for further development of the theory, in that it invites the use of already developed formalisms (e.g., Hamiltonian, Lagrangian, sum-over-histories) that have been applied in the context of the usual Schrödinger equation.[16]

Along with the preferred basis and trigger problems mentioned in Sec. I to which the GRW theory and the theory presented here respond, there are at least two other important problems for which there is as yet no response.

(3) *The relativity problem.* How can you make a relativistically invariant theory of state-vector reduction?

(4) *The link problem.* What is the connection of the reduction mechanism with the rest of physics? Is there some aspect of an already known field that is responsible for reduction?

There is an indication that this continuous spontaneous localization theory may be useful in solving problem (3). A number of authors[17–19] have described a conflict of relativity with quantum theory-plus-instantaneous state-vector reduction.

Consider the canonical two-packet situation, as seen from one Lorentz frame. Suppose that for $t<0$ the squared amplitude associated with each packet is $\frac{1}{2}$, but that for $t>0$ the squared amplitude of one packet is 1 and of the other is 0, i.e., this is the frame in which instantaneous state-vector reduction occurs. Then in other Lorentz frames there will be intervals of time over which the sum of squared amplitudes is either $\frac{1}{2}$ or $\frac{3}{2}$. that is, the norm of the state vector is not always equal to 1 in these frames, and of course this means that the usual quantum theory is not valid in these frames over these time intervals. But state-vector norms which are not equal to 1 are the *basis* of the theory presented here, so

one might hope for progress using these ideas.

Regarding problem (4), there have been a number of suggestions that gravity may be linked with state-vector reduction.[20–24] We close with the remark that a continuous spontaneous localization theory might solve a problem of semiclassical gravity.

Semiclassical gravity is a theory in which the classical Einstein equation of general relativity has as its source the quantum expectation value of the energy-momentum tensor. Kibble[25] has shown how to obtain the coupled Schrödinger equation and Einstein equation from an action principle, by varying the state vector and the metric tensor. However, he has also pointed out that semiclassical gravity has an obvious conflict with experiment.

Consider a quantum experiment for which the Schrödinger equation describes the evolution of a physical system into the canonical two-packet superposition, where the packets describe the center of mass position of a massive object such as a bowling ball. The expectation value of the energy-momentum tensor in such a state is the same as if there actually were two large masses located at the packet sites, and the metric tensor in the semiclassical theory responds to both masses. Of course, in an actual experiment (and a rather tongue-in-cheek experiment has actually been performed[26]), the mass is to be found centered on only one of the packet sites, and the metric tensor responds accordingly. We remark that, if the Einstein equation were coupled to a Schrödinger equation which included the terms given here that reduce the state vector, this embarrassing conflict of semiclassical gravity with experiment would disappear.

ACKNOWLEDGMENTS

Most of this work was performed at the Dipartimento di Fisica Teorica, Università di Trieste and the Dipartimento di Fisica Nucleare e Teorica, Università di Pavia. I am deeply indebted to GianCarlo Ghirardi and Alberto Rimini, whose generosity of time and ideas made this work possible. I am grateful to Hamilton College for support and to Professor Abdus Salam for hospitality at the International Centre for Theoretical Physics, Trieste. I also acknowledge financial support from the Istituto Nazionale di Fisica Nucleare (INFN), Sezioni di Pavia e Trieste, Italia.

APPENDIX A

We give here the *complete* probability density description of the evolution of an initial wave function $\psi(x,0)$ into a *physical* ensemble of *normalized* wave functions. First consider the behavior of squared amplitudes at M points uniformly spaced along the x axis, with spacing Δ. We define

$$Z_m(t)\equiv|\psi(x_m,t)|^2\sum_n\Delta|\psi(x_n,t)|^2=Y_n(t)/\Delta Y(t) \tag{A1}$$

($Y\equiv\sum_m Y_m$). $Z_m(t)$ approaches the squared amplitude of a normalized wave function in the continuum limit $M\to\infty$ followed by $\Delta\to 0$. However, even before the limit is taken

$$\sum_m \Delta Z_m=1\ . \tag{A2}$$

Our starting point is the Fokker-Planck Eq. (24a),

$$\partial\rho(\mathbf{Y};t)/\partial t=2\lambda\sum_{n,m}\Phi(x_n-x_m)(\partial/\partial Y_n)(\partial/\partial Y_m)Y_nY_m\rho\ , \tag{A3}$$

for the probability density $\rho(\mathbf{Y};t)$ of the unnormalized squared amplitudes $\mathbf{Y}_n$ in the raw ensemble. From this we will be able to obtain the Fokker-Planck equation for the probability density W of the Z_m's in the physical ensemble.

Equation (A2) states that there is a linear relationship between the Z_m's. In order to treat the Z_m's as independent variables we employ the following trick. We increase the number of independent variables to $M+1$, treating Δ^{-1} as an independent variable along with $\mathbf{Y}$, and Y as an independent variable along with $\mathbf{Z}$. We replace ρ by

$$\rho(\mathbf{Y};t)\to\rho(\mathbf{Y},\Delta^{-1};t)\equiv\rho(\mathbf{Y};t)\delta(\Delta^{-1}-\Delta_0^{-1})\ , \tag{A4}$$

which still satisfies Eq. (A3). The probability density of $\mathbf{Z}, Y$,

$$W(\mathbf{Z},Y;t)d\mathbf{Z}\,dY \equiv(\Delta\sum_n Y_n)\rho(\mathbf{Y};t)\delta(\Delta^{-1}-\Delta_0^{-1})d\mathbf{Y}\,d\Delta^{-1}\ , \tag{A5}$$

will be normalized to 1 with the choice $\Delta_0^{-1}=[\sum_n Y(x_n,0)]$ since

$$\int W\,d\mathbf{Z}\,dY=\Delta_0\sum_n\langle Y_n(t)\rangle=\Delta_0\sum_n Y_n(0)\ , \tag{A6}$$

the last step following from the Martingale nature of each Y_n. The $(\Delta\sum_n Y_n)$ factor in Eq. (A5) approaches the squared wave-function norm in the continuum limit, and therefore is the right factor to correct the raw ensemble to the physical in that limit.

Actually, we are most interested in the probability distribution of the Z_m's regardless of Y,

$$\begin{aligned}W(\mathbf{Z},t)&\equiv\int dY\,W(\mathbf{Z},Y;t)\\&=\int_0^\infty dY\,\Delta_0 Y\rho(\mathbf{Z}\Delta_0 Y;t)\delta(\sum_n Z_n-\Delta_0^{-1})\\&\quad\times J(\mathbf{Z}\Delta_0 Y,\Delta_0^{-1}|\mathbf{Z},Y)\ ,\end{aligned} \tag{A7}$$

obtained by integrating Eq. (A5) over Y, and by using Eqs. (A1) and (A2) to write $\mathbf{Y}$, Δ^{-1} in terms of $\mathbf{Z}, Y$. In Eq. (A7), $J(\mathbf{Y},\Delta^{-1}|\mathbf{Z},Y)$ is the Jacobian of the transformation from $(\mathbf{Y},\Delta^{-1})$ to $(\mathbf{Z},Y)$, and is readily calculated to be

$$J(\mathbf{Y},\Delta^{-1}|\mathbf{Z},Y)=(Y/\sum_n Z_n)^{M-1}\ . \tag{A8}$$

Now, to obtain the Fokker-Planck equation for W, take the derivative of Eq. (A7) with respect to t. Substitute Eq. (A3) for $\partial\rho/\partial t$ into the right-hand side of the resulting equation. Use

$$\partial/\partial Y_n=(S/Y)\partial/\partial Z_n-(1/Y)\sum_m Z_m\partial/\partial Z_m+\partial/\partial Y \tag{A9}$$

($S\equiv\sum_m Z_m$), which follows from Eq. (A1), to replace derivatives with respect to Y_n's by derivatives with

respect to Z_n's. After integration over Y, and after some manipulation we obtain

$$\partial W(\mathbf{Z},t)/\partial t = 2\lambda \sum_{n,m} \Phi(x_n - x_m) \left[S\frac{\partial}{\partial Z_n} - \sum_r \left[\frac{\partial}{\partial Z_r}\right] Z_r \right] \times \left[\frac{\partial}{\partial Z_n} - \sum_r \left[\frac{\partial}{\partial Z_r}\right] \frac{Z_r}{S} \right] \times Z_n Z_m W S^{-2} . \quad \text{(A10)}$$

After further manipulation, Eq. (A10) can be put in the forms

$$\partial W(\mathbf{Z},t)/\partial t = 2\lambda \sum_{n,m} (\partial/\partial Z_n)(\partial/\partial Z_m)\mu_{nm}(\mathbf{Z}) W \quad \text{(A11a)}$$

$$= \lambda \sum_{n,m} (\partial/\partial Z_n - \partial/\partial Z_m)^2 [-\mu_{nm}(\mathbf{Z})] W , \quad \text{(A11b)}$$

$$\mu_{nm}(\mathbf{Z}) \equiv Z_n Z_m \left[\Phi_{nm} + S^{-2} \sum_{j,k} \Phi_{jk} Z_j Z_k - S^{-1} \sum_k (\Phi_{nk} + \Phi_{mk}) Z_k \right] ,$$

$$\Phi_{nm} \equiv \Phi(x_n - x_m) . \quad \text{(A11c)}$$

Equation (A11) describes a gambler's ruin game among the Z_m's. Since any function of $S = \sum_k Z_k$ commutes with the differential operators $(\partial/\partial Z_n - \partial/\partial Z_m)$ in Eq. (A11b), if W is initially proportional to $\delta(\Delta S - 1)$ it remains so. Thus the amount the "money" in the "game" is preserved. Likewise, it follows from Eq. (A11) that $d\langle Z_n \rangle/dt = 0$, so each Z_n is a Martingale, and the game is "fair."

The complicated dependence of the diffusion constants on the Z_n's may be interpreted as a rule whereby each player's gain and rate of play depends upon the amount of money possessed by all. It is not hard to show that the probability distribution of Z_i/Z_j does not change with time if $|x_i - x_j| \ll \alpha^{-1/2}$. In other words, "nearby" gamblers win and lose together in such a way that the ratio of their wealth is constant. We remark that, in the continuum limit, Eq. (11b) can be written as a Fokker-Planck *functional* differential equation,

$$\partial W(\mathbf{Z}(x);t)/\partial t = 2\lambda \int\int dx\, dx' [\delta/\delta Z(x) - \delta/\delta Z(x')]^2 Z(x) Z(x') \times \left[\int dx_1 [\Phi(x - x_1) + \Phi(x' - x_2)] Z(x_1) - \Phi(x - x') - \int\int dx_1 dx_2 \Phi(x_1 - x_2) Z(x_1) Z(x_2) \right] W . \quad \text{(A11d)}$$

As an interesting application of the use of Eq. (11), one can calculate the equation of evolution of the density matrix,

$$D(X,X';t) \equiv \int d\mathbf{Z}\, W(\mathbf{Z};t) \psi(x,t) \psi^*(x',t) = e^{i\vartheta(x) - i\vartheta(x')} \times \int d\mathbf{Z}\, W(\mathbf{Z};t) [Z(x) Z(x')]^{1/2} \quad \text{(A12)}$$

$[\vartheta(x,t) \equiv (2i)^{-1} \ln \psi(x,t)/\psi^*(x,t)]$. Taking the derivative of Eq. (A12) with respect to t, substituting (A11) for $\partial W/\partial t$, and integrating by parts over $\mathbf{Z}$ yields

$$\partial D(x,x';t)/\partial t = \lambda [\Phi(x - x') - 1] D(x,x';t) , \quad \text{(A13)}$$

which was obtained much more easily in Sec. VII, Eq. (39).

It is easy to apply Eq. (A11b) to the situation of K wave packets, each of width $\ll \alpha^{-1/2}$ separated by distances $\gg \alpha^{-1/2}$. We define the squared norm of the kth packet as

$$X_k(t) \equiv {}_k \sum_m \Delta Z_m \to {}_k \int dx\, Z(x,t) , \quad \text{(A14)}$$

where the subscript k restricts the sum or integral to the x interval of support of the kth packet. The arrow refers to the continuum limit where $Z_m \to Z(x_m,t)$ is the square of the normalized wave function at x_m.

To find the probability density R of the packet squared norms

$$R(\mathbf{X};t) \equiv \int d\mathbf{Z}\, W(\mathbf{Z};t) \prod_{k=1} \delta({}_k \sum_n \Delta Z_n - X_k) , \quad \text{(A15)}$$

we take the derivative of Eq. (A15) with respect to t, and substitute Eq. (A11b) for $\partial W(\mathbf{Z},t)/\partial t$. The diffusion coefficients (A11c) in this situation are

$$\mu_{nm} = Z_n Z_m \left[\delta_{k(n)k(m)} + \sum_k X_k^2 - X_{k(n)} - X_{k(m)} \right] , \quad \text{(A16)}$$

where we have used $\Phi_{nm} \approx 1$ or 0 depending upon whether x_n and x_m are in the same packet or not. In Eq. (A16), $k(n)$ denotes the index of the packet which has support at the point x_n. After integrating by parts, and converting the derivatives with respect to the Z_n's to derivatives with respect to the X_k's by means of the δ functions in Eq. (A15), the following result is obtained:

$$\frac{\partial R(\mathbf{X};t)}{\partial t} = 2\lambda \sum_{j,k} \left[\frac{\partial}{\partial X_j} - \frac{\partial}{\partial X_k} \right]^2 X_j X_k \times \left[X_j + X_k - \sum_i X_i^2 \right] R(\mathbf{X};t) . \quad \text{(A17)}$$

Equation (A17) describes a gambler's ruin game among the squared packet norms X_k. It's significance is discussed in Sec. VI.

APPENDIX B

In this Appendix we consider the canonical two-packet situation in one spatial dimension *in the GRW theory*. We obtain an expression for the probability density distribution $R(X_1,t)$, where

$$X_k(t)\equiv\int_k dx\,|\psi(x,t)|^2 \quad (k=1 \text{ or } 2) \tag{B1}$$

is the squared norm associated with packet k belonging to the normalized wave function ψ (so $X_2=1-X_1$).

Consider an ensemble described by $R(X_1,t)$. According to GRW, in the time interval dt there is probability $\lambda\,dt$ that a particular wave function will be hit. If ψ is hit with a center at z, the wave function immediately after the hit is

$$(\alpha/\pi)^{1/4}\{\exp[-(\alpha/2)(x-z)^2]\}\psi(x,t)/N(z,t)\ , \tag{B2}$$

where $N(z,t)$ normalizes the new wave function and $N^2(z,t)dz$ is the probability that the center lies between z and $z+dz$,

$$N^2(z,t)\equiv(\alpha/\pi)^{1/2}\int dx\,\{\exp[-\alpha(x-z)^2]\}\,|\psi(x,t)|^2 \tag{B3a}$$

$$\approx(\alpha/\pi)^{1/2}(X_1\{\exp[-\alpha(x_1-z)^2]\}+X_2\{\exp[-\alpha(x_2-z)^2]\})\ . \tag{B3b}$$

In obtaining Eq. (B3b) from (B3a) we have used the approximation that packets 1 and 2 are narrowly spread about points x_1 and x_2, respectively. In the same approximation, by squaring Eq. (B2) and inserting it into (B1), we obtain the result that a packet characterized by $X_1=X'$ before a hit becomes characterized by

$$X=a_1X'/[a_1X'+a_2(1-X')] \tag{B4a}$$

($[a_k\equiv\exp[-\alpha(x_k-z)^2]$) after the hit. The inversion of Eq. (B4a) is

$$X'=a_2X/[a_2X+a_1(1-X)]\ . \tag{B4b}$$

The probability $R(X;t+dt)dX$ that X lies in the range dX at time $t+dt$ has two contributions. The probability that there is a transition into this range is, using (B3b),

$$\lambda\,dt(\alpha/\pi)^{1/2}\int dz[a_1X'+a_2(1-X')]R(X';t)dX'|_z\ , \tag{B5a}$$

where it is understood that X' in Eq. (B5a) is the expression (B4b), which is a function of z and X. The probability that there is a transition out of this range is

$$\lambda\,dt(\alpha/\pi)^{1/2}\int dz[a_1X+a_2(1-X)]=\lambda\,dt\ . \tag{B5b}$$

Therefore the equation of evolution for $R(X;t)$ is

$$\partial R(X;t)/\partial t=\lambda(\alpha/\pi)^{1/2}\int_{-\infty}^{\infty}dz[a_1X'+a_2(1-X')]R(X';t)(\partial X'/\partial X)_z-\lambda R(X;t)\ . \tag{B6}$$

In order to solve Eq. (B6), it is useful to change the variable of integration from z to X'. We note from Eq. (B4b) that, as z ranges from $-\infty$ to $+\infty$, X' ranges from 0 to 1 (with the choice $x_2\geq x_1$). By use of $(\partial z/\partial X')_X(\partial X'/\partial X)_z=-(\partial z/\partial X)_{X'}$ Eq. (B6) becomes

$$\partial R(X;t)/\partial t=\lambda(\alpha/\pi)^{1/2}\int_0^1 dX'\{[a_1X'+a_2(1-X')]^3/2\alpha(x_2-x_1)a_1a_2X'(1-X')\}R(X';t)-\lambda R(X;t) \tag{B7a}$$

$$=\lambda[(\alpha/\pi)^{1/2}/2\alpha(x_2-x_1)X^2(1-X)]\int_0^1 dX'X'a_1[z(X,X')]R(X';t)-\lambda R(X;t)\ . \tag{B7b}$$

By solving Eq. (B4) for z in terms of X,X' we obtain

$$a_1[z(X,X')]=[X(1-X')/X'(1-X)]^{1/2}\exp[-(\alpha/4)]((x_2-x_1)^2+(x_2-x_1)^{-2}\{\ln[X(1-X')/X'(1-X)]\}^2)\ . \tag{B8}$$

Remarkably, it turns out that $X'a_1$ multiplied by the bracketed expression outside the integral in Eq. (B7b) is $G(X,X';\alpha(x_2-x_1)^2)$, where $G(X,X';T)$ is the solution Eq. (35b) of Eq. (36),

$$\partial G(X,X';T)/\partial t=(\partial/\partial X)^2[X^2(1-X)^2G(X,X';T)]\ , \tag{B9}$$

with the initial condition $G(X,X';0)=\delta(X-X')$. Thus our equation for R is

$$\partial R(X;t)/\partial t=\lambda\int_0 dX'G(X,X';\alpha(x_2-x_1)^2)R(X';t)-\lambda R(X,t)\ . \tag{B10}$$

We can now proceed to solve Eq. (B10). The explicit expression for $G(X,X';T)$ appears in Eq. (35b) (with $4\lambda\to1$, $t\to T$, $X_1\to X$, and $X_{10}\to X'$). Since $\alpha(x_2-x_1)^2\gg1$, we only need to know G for large $T=\alpha(x_2-x_1)^2$,

$$G(X,X';T)\approx X'\delta(1-X)+(1-X')\delta(X)\ . \tag{B11}$$

The δ functions in Eq. (B11) are actually narrow peaks of height $\sim e^T$ and of area close to 1, centered at $X=\{1+[e^{\pm T}(1-X')(X')^{-1}]\}^{-1}$. Putting the expression (B11) for G into Eq. (B10) yields

$$\partial R(X;t)/\partial t=\lambda[X'R(1;t)+(1-X')R(0;t)-R(X;t)] . \tag{B12}$$

The solution of Eq. (B12), subject to the initial condition $R(X;0)=\delta(X-X_0)$ is

$$R(X;t)=\delta(X-X_0)\exp(-\lambda t) +[X_0\delta(1-X)+(1-X_0)\delta(X)] \times[1-\exp(-\lambda t)] . \tag{B13}$$

The significance of this solution is discussed in Sec. VII.

APPENDIX C

In this Appendix we explain how the probability weighting of the physical ensemble is consistent with an independent Markovian evolution for each state vector in the ensemble. Actually, the truth of this assertion follows immediately from Eq. (A11), which shows that the probability density of the normalized state vectors in the physical ensemble obeys a Fokker-Planck equation. However, here we wish to provide a more intuitive understanding.

Consider an ensemble of state vectors, each of which evolves for $0\leq t<T$, according to the Schrödinger equation, by means of one of a set of possible non-Hermitian Hamiltonians H_j, $j=1,\ldots,J$. Suppose that a particular state vector "chooses" the Hamiltonian H_j according to a certain probability rule. Furthermore, suppose that for each successive time interval of T seconds a (possibly) new Hamiltonian is similarly chosen.

The probability P_j that the jth Hamiltonian is chosen might depend upon the time interval, upon the history and present state of the state vector making the choice, and indeed upon the histories and present states of all the other state vectors in the ensemble. However, it is natural to choose the probabilities P_j to be fixed numbers (an especially simple choice is $P_j=1/J$) because of the following beneficial consequences.

It is a rudimentary *time-translational invariant* process because the probability rule does not depend upon the time interval. (In the limit $T\rightarrow 0$, it is truly time translational invariant.)

It is a rudimentary *Markovian* process because the probability rule for a given interval does not depend upon the history of any state vector in the ensemble in any previous interval, and the Hamiltonian evolution depends only on the choice made at the beginning of the interval. (In the limit $T\rightarrow 0$, it is truly Markovian.)

Each state vector *evolves independently* because the probability of the choice made by a particular state vector depends only upon the Hamiltonian to be chosen (i.e., upon the index j), and not upon the other state vectors in the ensemble. In fact, the probability does not even depend upon that state vector itself, and here is the point.

The probability rule can be generalized without losing these three properties. If the rule depends upon the state vector making the choice as well as upon the index j, that state vector still evolves independently. If in addition the rule is independent of the time interval, and depends only upon the state vector at the beginning of the time interval, we have all three properties. Now we will choose a peculiar probability rule that satisfies these requirements, but it is a rule that is only suitable for a set of Hamiltonians H_j with a peculiar property.

Consider a particular state vector which has evolved to the beginning of a particular time interval. Generally that state vector's norm will not equal 1. So, normalize it. Now, consider the evolution $|\psi\rangle\rightarrow V_j(T)|\psi\rangle$ of that normalized state vector over the time interval T by means of the Hamiltonian H_j, i.e., $i\,dV_j(t)/dt=H_jV_j(t)$. The state vector's squared norm at time T is $N_j^2=\langle\psi|V_j(T)^*V_j(T)|\psi\rangle$. Our rule is $P_j=N_j^2/J$.

Note that N_j^2, and therefore P_j, depends only upon the state vector at the beginning of the time interval and upon the chosen Hamiltonian H_j. Thus we have the three desirable properties.

This rule only makes sense if the probabilities sum to 1 for an arbitrary state vector $|\psi\rangle$,

$$\sum_j P_j=J^{-1}\sum_j N_j^2\equiv\langle N^2\rangle=1 \quad \text{for any } |\psi\rangle , \tag{C1a}$$

so this is a necessary consistency condition for the probability rule [see Eq. (9)]. It can be achieved for a set of Hamiltonians $\{H_j\}$ if and only if

$$J^{-1}\sum_j V_j(T)^*V_j(T)=1 \tag{C1b}$$

and we will suppose this is the case.

Finally, we can obtain an interesting result concerning the norms of an ensemble of state vectors at time $t=NT$ which evolved from a single state vector $|\psi,0\rangle$.

Consider one of the evolved state vectors $|\psi_\Omega,NT\rangle\equiv|\psi,j(N),\ldots,j(1)\rangle$ in the ensemble which utilized the $j(n)$th Hamiltonian at the nth time interval. The probability that it is in the ensemble is

$$P(|\psi_\Omega,NT\rangle)=\prod_{n=1}^{N}P_{j(n)} =J^{-N}\prod_{n=1}^{N}N_{j(n)}^2[j(n),\ldots,j(1);|\psi,0\rangle] . \tag{C2}$$

[We have indicated in Eq. (C2) that the squared norm $N_{j(n)}^2$ acquired during the nth interval depends not only on $j(n)$, but also on the state vector at the beginning of the nth interval, and that state vector in turn depends upon the initial state vector and its subsequent history, up to that interval.]

On the other hand, the norm of this state vector is

$$\langle\psi,j(n),\ldots,j(1)|\psi,j(n),\ldots,j(1)\rangle=\langle\psi,0|V_{j(1)}(T)^*\cdots V_{j(N)}(T)^*V_{j(N)}(T)\cdots V_{j(1)}(T)|\psi,0\rangle =\prod_{n=1}^{N}N_{j(n)}^2[j(n),\ldots,j(1);|\psi,0\rangle] . \tag{C3}$$

Combining Eqs. (C2) and (C3) we obtain the result

$$P(|\psi_\Omega,NT\rangle)=J^{-N}\langle\psi_\Omega,NT|\psi_\Omega,NT\rangle\ . \tag{C4}$$

Thus the probability that the state vector $|\psi_\Omega,NT\rangle$ belongs to the ensemble at time NT is proportional to the norm of that state vector. The proportionality factor J^{-N} is the probability that a particular "path" of Hamiltonians is chosen according to the simple rule that all Hamiltonians are equally likely.

It is to be expected, with an appropriate choice of Hamiltonians, that an appropriate limit $T\to 0$ can be taken in which the process described in this Appendix approaches the process described in the main body of this paper. In particular, Eq. (C4) becomes the probability rule $P_\Omega = d\Omega\, N_\Omega^2$ of the physical ensemble.

---

[1]D. Bohm and J. Bub, Rev. Mod. Phys. **38**, 453 (1966).

[2]P. Pearle, Phys. Rev. D **13**, 857 (1976). For recent reviews see *Quantum Concepts in Space and Time,* edited by R. Penrose and C. J. Isham (Clarendon, Oxford, 1986), p. 84; Ann. N.Y. Acad. Sci. **480**, 539 (1986).

[3]N. Gisin, Phys. Rev. Lett. **52**, 1657 (1984). See also P. Pearle, Phys. Rev. Lett **53**, 1775 (1984); N. Gisin, Phys. Rev. Lett. **53**, 1776 (1984); L. Diosi, J. Phys. A: Math. Nucl. Gen. **21**, 2885 (1988).

[4]P. Pearle, Found. Phys. **12**, 249 (1982).

[5]W. Feller, *An Introduction to Probability Theory and its Applications* (Wiley, New York, 1950), Chap. 14.

[6]G. C. Ghirardi, A. Rimini, and T. Weber, Phys. Rev. D **34**, 470 (1986). For a recent review see Found. Phys. **18**, 1 (1988).

[7]J. S. Bell, in *Schrodinger-Centenary Celebration of a Polymath,* edited by C. W. Kilmister (Cambridge University Press, Cambridge, 1987), p. 41.

[8]In the GRW theory and in the theory presented here the state vector is never completely reduced. There is always a small but nonvanishing piece of "what might have been" included in the state vector. We do not regard this as satisfactory. If the reduced state vector is to correspond to what is actually observed in nature, it is hard to see what meaning can be given to an additional term that describes another observation, no matter how small its coefficient may be. (We have fulminated over this issue in Refs. 3 and 12, pointing out that the theory described in Ref. 2 does not have this defect.) I hope that further development of the theory presented here will contain correction terms leading to complete reduction, perhaps because it may possess a more realistic noise source than white noise.

[9]Since $\langle \int dx\, f(x)dw(x,t)\int dx' f(x')dw(x',t)\rangle > 0$ for arbitrary $f\neq 0$, Φ must be positive definite, i.e., $\int\int dx\, dx' f(x)\Phi(x-x')f(x')>0$. Any such function will do which also has $\Phi(0)=1$, $\Phi(\infty)=0$, and a characteristic length governing the transition of Φ from 1 to 0.

[10]Taking $H=0$ amounts to the assumption that the reduction dynamics takes place over a time interval short compared to the time characterizing the Hamiltonian evolution. This will often not be the case. However, we will not discuss in this paper the interference between the Hamiltonian and reduction evolutions. In the canonical two-packet situation upon which we focus our examples, inclusion of the Hamiltonian is an inessential complication: a wave function consisting of two widely separated moving packets responds to the reduction process in the same way as if the packets were at rest.

[11]More precisely, if we define $\vartheta(x,t)\equiv(2i)^{-1}\ln[\psi(x,t)/\psi^*(x,t)]$, we calculate $\langle d\vartheta\rangle=0$, $\langle(d\vartheta)^2\rangle=0$ using Eqs. (4) and (5) and Eqs. (15) and (26). Since $\vartheta(x,t)$ has vanishing drift and diffusion, it does not change with time.

[12]P. Pearle, Phys. Rev. D **33**, 2240 (1986).

[13]It has, however, been recently accomplished [G. C. Ghirardi, O. Nicrosini, A. Rimini and T. Weber (unpublished).

[14]G. C. Ghirardi, P. Pearle and A. Rimini (unpublished).

[15]P. Pearle, Phys. Rev. D **29**, 235 (1984).

[16]In this regard it is worth pointing out that the Stratonovich equation, equivalent to the Itô equation (4) or (51) upon which this theory is based, is also linear in ψ. For example, Eq. (4) is $\partial\psi(x,t)/\partial t=-iH\psi(x,t)dt+[\partial w(x,t)/\partial t-\lambda]\psi(x,t)$ in Stratonovich form [i.e., the $\frac{1}{2}$ in Eq. (4) is replaced by 1]. The Stratonovich form may be manipulated using the rules of ordinary calculus, and it is this form that arises when one considers more realistic noise sources for $\partial w(x,t)/\partial t$ than pure white noise [N. Van Kampen, Phys. Rep. **24**, 171 (1976)]. It can be shown [E. Wong and M. Zakai, Ann. Math. Stat. **36**, 1560 (1965)] that, if a sequence of functions converging toward white noise is inserted in a Stratonovich equation (making it a standard partial differential equation with time-dependent coefficients), the behavior of the limit of the sequence of the corresponding solutions of the equation is described by the Fokker-Planck equation (23).

[17]I. Bloch, Phys. Rev. **156**, 1377 (1967).

[18]Y. Aharanov and D. Z. Albert, Phys. Rev. D **21** , 3316 (1980); **24**, 359 (1981).

[19]S. Malin, Phys. Rev. D **26**, 1330 (1982).

[20]F. Karolyhazy, A. Frenkel, and B. Lukacz, in *Physics as Natural Philosophy,* edited by A. Shimony and H. Feshbach (M.I.T. Press, Cambridge, MA, 1982), P. 204; *Quantum Concepts in Space and Time,* edited by R. Penrose and C. J. Isham (Clarendon, Oxford, 1986), p. 109.

[21]P. Pearle, Int. J. Theor. Phys. **18**, 489 (1979).

[22]R. Penrose, in *Quantum Gravity 2, a Second Oxford Symposium,* edited by C. J. Isham, R. Penrose, and D. W. Sciama (Oxford University Press, Oxford, 1981); *Quantum Concepts in Space and Time,* edited by R. Penrose and C. J. Isham (Clarendon, Oxford, 1986).

[23]D. W. Sciama (private communication).

[24]L. Diosi (unpublished).

[25]T. W. Kibble, in *Quantum Gravity 2, a Second Oxford Symposium,* edited by C. J. Isham, R. Penrose, and D. W. Sciama (Oxford University Press, Oxford, 1981), p. 63.

[26]D. N. Page and C. D. Geilker, Phys. Rev. Lett. **47**, 979 (1981).