INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

COMBINING STOCHASTIC DYNAMICAL STATEVECTOR REDUCTION WITH SPONTANEOUS LOCALIZATION

Philip Pearle

INTERNATIONAL ATOMIC ENERGY AGENCY

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

1988 MIRAMARE-TRIESTE
1. Dynamical Reduction

Dynamical reduction of the statevector is a process whereby a superposition such as

$$|ψ,0⟩ = a_{1}(0) |ψ_{1}⟩ + a_{2}(0) |ψ_{2}⟩$$

continuously evolves with time into

$$|ψ,t⟩ = e^{iθ_{k}} |ψ_{k}⟩ \quad (k = 1 \text{ or } 2)$$

A number of theories have so far been constructed\(^1,2,3\) which modify the usual Schrödinger equation to achieve dynamical reduction. They enable one to interpret the statevector 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 $X_{k}(t) = |a_{k}(t)|^2$ to fluctuate randomly, although their sum $\sum X_{k}(t) = 1$ 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 statevectors 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

$$d<X_{k}(t)>/dt = 0$$

\(^1\) Submitted for publication.

\(^2\) Most of this work was performed at the Dipartimento di Fisica Teorica, Università di Trieste, Trieste, and the Dipartimento di Fisica Nucleare e Teorica, Università di Pavia, Pavia, Italy.

\(^3\) Present address until 20 June, 1988; Dipartimento di Fisica Nucleare e Teorica, Università di Pavia, Pavia, Italy.
(the average <> is over the ensemble of trajectories of $X_k(t)$) that ensures agreement with the predictions of quantum theory: the outcome $X_k(t) \rightarrow 1$ in Eq. (2) occurs with probability $P_k = X_k(0)$. Here is the argument. From Eq. (3), $<X_k(0)> = <X_k(t)>$. On the one hand, $<X_k(0)> = X_k(0)$ since we start with the pure state (1). On the other hand, for sufficiently large $t$ so that Eq. (2) obtains for all statevectors in the ensemble, $<X_k(t)> = 0 \cdot \text{(probability } X_k \rightarrow 0) + 1 \cdot \text{(probability } X_k \rightarrow 1)$. Thus, $X_k(0) = P_k$.

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

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. $2^{-1/2}(\mid k_1\rangle \pm \mid k_2\rangle)$)?

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?

### 2. Spontaneous Localization

Recently, Ghirardi, Rimini and Weber⁶ (see also Bell⁷) 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 statevector 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 statevector should be after a hit. If the wavefunction of a single particle before the hit is $\psi(x)$, after the hit it is multiplied by a Gaussian $\exp(-\alpha/2)(x-x)^2$, and divided by a numerical factor $N$ to correctly normalize the altered wavefunction. (GRW have chosen the localization width $\alpha^{-1/2} = 10^{-5}$ cm.) We shall call the point $x$ 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 $\lambda = 10^{-16}$ sec$^{-1}$ = $10^{-8}$ year$^{-1}$). Thus centers are most likely to appear where the wavefunction is largest.

A wavefunction well localized within $\alpha^{-1/2}$ will scarcely be affected by a hit. However, a wavefunction 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 wavefunction. 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 wavefunction. The frequency of such a hit is proportional to the number of particles in the body, so such a reduction will take place rapidly.

### 3. 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 statevector. 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 wavefunction:

\[ \text{\textit{d}v(x,t) = -iHv(x,t)dt + [dw(x,t) - (1/2)\lambda(t)]v(x,t)} \]  
(4)

Eq. (4) is remarkable, for a dynamical reduction theory, in that it is linear in \( \psi \). \( H \) 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:

\[ \langle dw(x,t) \rangle = 0 \]  
(5a)

\[ \langle dw(x,t)dw(x,t') \rangle = \lambda \delta(t-t') \]  
(5b)

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

The function \( \Phi \) suggested by the GRW theory is

\[ \Phi_{\text{GRW}}(x-x') = \exp(-a/(x-x')^2) \]  
(6)

but there are other suitable choices. 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:

\[ \langle dB(z,t) \rangle = 0 \]  
(7a)

\[ \langle dB(z,t)dB(z',t) \rangle = \delta(z-z')\lambda dt \]  
(7b)

(Eqs. (5), (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 sections 4,5,6. We will assume for simplicity that \( H=0 \), in order to concentrate on the effect of the new terms in Eq. (4). Their most obvious feature is that they change the norm of the wavefunction.

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

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

It might appear that this precept makes the probability that a particular statevector is in the ensemble dependent upon the composition of the rest of the ensemble. This would be unacceptable, because an ensemble of independently evolving statevectors is an ensemble described by classical probability theory, in which there can be no interference of probabilities. That this precept is, in fact, consistent with an independent Markovian evolution of each statevector 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 \( d\Omega \) be the probability measure in the space of Brownian functions, and \( w_\Omega \) be the particular Brownian function responsible for the evolution of the particular wavefunction \( \psi_\Omega \), with squared norm

\[ N_{\psi_\Omega}^2(t) = \int dx |\psi_\Omega(x,t)|^2 \]  
(8)

Then, according to our precept, the probability that \( \psi_\Omega \) lies in the ensemble is \( N_{\psi_\Omega}^2 d\Omega \), not \( d\Omega \). Of course this interpretation requires, for consistency, that the sum of probabilities remains equal to 1

\[ \int N_{\psi_\Omega}^2(t) d\Omega = \int N_{\psi_\Omega}^2(t) = 1 \]  
(9)

which we will see (Eq. (18a)) follows from Eq. (4). We shall refer to the ensembles with weights \( d\Omega \) and \( N_{\psi_\Omega}^2 d\Omega \) as the raw and the physical ensembles respectively.

Now, what will be the behavior of the ensemble of norms, and which wavefunctions will predominate because their norms are largest?
Let us write out the essential part of Eq. (4) as

\[ y(x,t+\Delta t) = (1 + \Delta w(x,t) - (1/2)\Delta \beta(t)) y(x,t) \]

- \[ y(x,t+\Delta t) = (1 + \int dz \Delta B(z,t) \exp(-z^2/(2\Delta \beta(t))) y(x,t) \]  

One expects that, since \( \Delta w \) (or \( \Delta B \)) 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 wavefunctions evolving subject to these Brownian motions will decrease roughly exponentially due to the last term in the bracket of Eq. (10). Thus these wavefunctions will be essentially unaffected in shape, but after a few time-constants they will be of negligible norm and of negligible importance in the physical ensemble.

On the other hand, for that minority of Brownian motions for which \( \Delta w(x,t) \), or a set of \( \Delta B(z,t) \) with \( x-\alpha \leq z \leq x+\alpha \), happens to be of one sign significantly more frequently than the other sign, the norm of the associated wavefunctions will grow if \( x \) also happens to have an \( \alpha^{-1/2} \) - neighborhood where \( y(x) \) 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 \( \Delta B(z,t) \) will increase the amplitude of the wavefunction in the neighborhood of \( x-z \), at the expense of the amplitude of the wavefunction 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 \( \Delta w \) 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 \( \Delta w \) 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 \( y \) should grow in the \( \alpha^{-1/2} \) - neighborhood of \( x \), it is most likely that there will be a concomitant approximate exponential decrease of \( y \) outside that neighborhood.

After a few time constants, the ensemble that results from a wavefunction 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" wavefunction hereafter), is as follows. In the raw ensemble, the huge majority of wavefunctions still contain two packets, but their norms have decreased exponentially. There is also a small set of wavefunctions 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 wavefunctions containing two packets is negligibly small. The overwhelmingly probable wavefunctions in the ensemble consist of essentially just one packet.

Now, what does the theory predict for the probability associated with the wavefunctions containing a particular 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 section 1 (e.g. \( |1 \rangle \), \( |2 \rangle \) in Eqs. (1), (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 section 1.

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

\[ d<|\psi(x,t)|^2>dt = 0 \]  

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 \( \psi_{\Omega}(x,t) \):

\[ N_{\Omega}(t) = \int dx |\psi_{\Omega}(x,t)|^2 \]  

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

\[ d<N_{\Omega}(t)>dt = 0 \]  

Now, the squared norm associated with \( \psi_{\Omega} \) is \( N_{\Omega}(t) = N_{\Omega_1}(t) + N_{\Omega_2}(t) \) (Eq. (8)). However, for \( t-T \) greater than a few time constants, either \( N_{\Omega_1} \) or \( N_{\Omega_2} \) or both essentially vanish as we have discussed. Therefore the probability \( P_k \), in the physical ensemble, that a wavefunction consists of just the \( k \)th packet, for \( t \geq T \), is
\[ P_k = \text{i} k N^2 \Omega(T) d\Omega = -N^2 \Omega(T) \]  
(14)

It follows from Eq. (13) that \( -N^2 \Omega(0) = -N^2 \Omega(T) \). But since \( -N^2 \Omega(0) = X_k(0) \), we obtain from Eq. (14) the quantum theory prediction \( P_k = X_k(0) \).

4. 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 \( \psi^*(x',t) \psi^2(x,t) \) behave in the raw ensemble (all we consider in this section) we utilize the Ito formula

\[ df = Xg + df + <dfdg> \]  
(15)

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

\[ d[\psi(x,t) \psi^*(x',t)] = \{H(\psi(x,t)) - H(\psi(x',t))\} dt + d\psi(x,t) + d\psi(x',t) + \lambda [\delta(x-x') - 1] \psi(x,t) \psi^*(x',t) \]  
(16)

(which will be useful later, when we discuss the density matrix in section 7) and, setting \( \lambda = H = 0 \)

\[ d\psi(x,t) = 2d\psi(x,t) \psi(x,t) \]  
(17)

Eqs. (17) and (5) contain the complete information about the ensemble of wavefunctions.

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

\[ <d\psi(x,t)> = 0, \quad <d\psi(x,t)>^2 = 4\Omega(\psi(x,t))^2 dt \]  
(18a, b)

(Note that Eq. (11) follows from Eq. (18a) while Eq. (9) follows from integrating Eq. (11) over \( x \).)

Using Eqs. (18a, b) we can construct the Fokker - Planck equation describing the behavior of the ensemble of squared amplitudes \( \psi(x) \) by means of the prescription

\[ \frac{\partial p(Y,t)}{\partial t} = -\frac{\partial}{\partial Y} \left[ <d\psi(x)> dt \right] p + (1/2) \left( \frac{\partial}{\partial Y} Y^2 \right) \left[ <d\psi(x)>^2 dt \right] p \]  
(19)

\( p(Y,t) dY \) is the probability that \( Y \) lies in the range \( dY \) in the raw ensemble. Putting Eqs. (18) into Eq. (19) we obtain

\[ \frac{\partial p(Y,t)}{\partial t} = 2\lambda (\partial^2 Y)^2 \{Y^2 p(Y,t)\} \]  
(20)

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

\[ p(Y,t) dY = \lambda (\partial Y)^2 \left[ \exp \left[ -\ln Y - \ln Y_0 + 2\lambda t \right] \right] dY \]  
(21)

Eq. (21) shows that \( p(0,t) = 0 \), and \( p(Y,t) \) rises to a peak at \( Y = Y_0 \exp -2\lambda t \), falling to zero as \( Y \rightarrow \infty \). Thus, in the raw ensemble, most wavefunctions have squared amplitudes which decrease roughly exponentially with time at each point.

Next consider the correlated behavior of \( \psi(x_1,t) \), \( \psi(x_2,t) \) at two different points \( x_1, x_2 \).

From Eq. (17) we find

\[ <d\psi(x_1,t) \psi(x_2,t)> = 4\lambda \delta(x_1-x_2) Y(x_1,t) Y(x_2,t) dt \]  
(22)

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

\[ \frac{\partial p(Y_1,Y_2,t)}{\partial t} = -\sum_{nm} \left[ (\partial Y_1) (\partial Y_2) \right] p + \frac{1}{2} \sum_{nm} \left[ (\partial Y_1) (\partial Y_2) \right] p (\partial Y_1) (\partial Y_2) p \]  
(23)

we obtain, for an arbitrary number of points

\[ \frac{\partial p(Y_{nm},t)}{\partial t} = 2\lambda \sum_{nm} \Phi(x_n-x_m) \delta(Y_{nm}) \left[ \partial Y_n \partial Y_m \right] p \]  
(24a)

which, for two points, is

\[ \frac{\partial p(Y_1,Y_2,t)}{\partial t} = 2\lambda \left[ (\partial Y_1)^2 (\partial Y_2)^2 + (\partial Y_1) (\partial Y_2)^2 \right] p + 2\delta(x_1-x_2) (\partial Y_1) (\partial Y_2) Y_1 Y_2 p \]  
(24b)

We see from Eq. (24b), when \( x_1 \) is distant from \( x_2 \) (i.e. \( |x_1-x_2| >> \alpha^{-1/2} \) so \( \phi = 0 \)), the evolutions of \( Y_1 \) and \( Y_2 \) are uncorrelated, and the raw probability distribution is the product of separate distributions (21) for \( Y_1 \) and \( Y_2 \). On the other hand, for nearby points \( |x_1-x_2| << \alpha^{-1/2} \) so \( \phi = 1 \), the solution of Eq. (24b) subject to the initial condition

\[ p(Y_1,Y_2,0) = \delta(Y_1-Y_1) \delta(Y_2-Y_2) \]  

\[ p(dY_1 dY_2) = \delta[(\ln Y_1)/Y_2] - \ln (Y_1/Y_2) [32x_1x_2]^{-1/2} \]  

\[ \exp[-(32x_1x_2)^{-1/2} [(\ln Y_1)/(Y_2) - (\ln Y_2)/(Y_2) + 4x_1x_2] dY_1 dY_2 \]  
(25)

Because of the \( \delta \)-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 wavefunction 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 $\Phi = 1$ employed above, we write down the equation of motion for the ratio $Y_1/Y_2$. Using Eq. (15) with $f = Y_1$ and $g = Y_2$, together with the Fokker-Planck formula

$$dQ(Y) = g'dY + (1/2)g''dY^2$$

with $g = 1/Y_2$ and Eq. (17), we obtain

$$d(Y_1/Y_2)^2 = [2dw(x_1,t) - dw(x_2,t)] + 4[1 - \exp(-ax/4)](x_1 - x_2)^2dtdY_1/Y_2$$

(26)

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

$$<d(Y_1/Y_2)>/dt = 4a[1 - \exp(-ax/4)](x_1 - x_2)^2(Y_1/Y_2)^2$$

(28)

$$<d(Y_1/Y_2)^2>/dt = 4<dw(x_1,t) - dw(x_2,t)^2>/dt$$

(29)

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

5. One Particle: Behavior of Packets

Now we turn to the canonical situation of an initial wavefunction consisting of two 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

$$dN_k(t)^2 = d\int(dx,Y) = 2\int dxdw(x,t)Y(x,t)$$

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

$$<dN_1^2>/dt = 0 , <dN_2^2>/dt = 0$$

(31a)

$$<dN_1^2dN_2^2> = 4\lambda \int dxdy \int dxdy \Phi(x-y)Y(x,t)Y'(y,t)dtdtdY_1/Y_2$$

(31b)

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

$$<dN_1^2dN_2^2> = 4\lambda(N_k^2)^2dtdY_1/Y_2 , <dN_1^2dN_2^2> = 0$$

(31c)

Thus the squared norms of the two packets are uncorrelated. In fact, by Eqs. (31a,c), the Fokker-Planck equation and its solution describing the distribution of $N_k^2$ in the raw ensemble is given by Eqs. (20), (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 $p(N_1^2,N_2^2;t)$ is given by

$$p(N_1^2,N_2^2;t)dN_1^2dN_2^2 = p(N_1^2,N_2^2)dN_1^2dN_2^2 = (8\pi)^{1/2}N_1^2dN_2^2 \cdot \exp(-8(N_1^2/N_2^2)^2 + 2N_2^2 + 2N_1^2/N_2^2 + 2N_2^2)$$

(32a)

(32b)

Eqs. (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 which the peak moves toward $N_k^2 = \delta(N_k^2 - N_{k0}^2)$ while the other peak moves toward $N_{k0}^2 = 0$. Moreover, the overall probability associated with each term is $N_{k0}^2\exp(-8N_{k0}^2/N_{k0}^2)\delta(N_k^2 - N_{k0}^2)$, the squared amplitude for the packet in the original wavefunction. Thus Eq. (32b) shows explicitly the reduction behavior, how the initial wavefunction with probability density $p(N_1^2,N_2^2;0) = \delta(N_1^2 - N_{10}^2)\delta(N_2^2 - N_{20}^2)$ evolves into an ensemble, with correct probability distribution, of wavefunctions of ever-increasing norm, each containing (essentially) just one packet.

6. Gambler's Ruin

In the previous sections we discussed the unnormalized wavefunctions. 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 wavefunctions, using the correct probability densities of the physical ensemble.

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

\[
Z(x,t) = \int |\psi(x,t)|^2 \, dx \| \psi(x,t) |^2
\]

is obtained in Appendix A, Eq. (A.11). 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 situation. We shall use the result (32b) of the previous section to find the probability density distribution for the squared norm of packet 1

\[
X_1(t) = \frac{1}{N_1^2(t)} \int (N_1^2(t) + N_2^2(t))
\]

(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 \( N^2 = N_1^2 + N_2^2 \):

\[
R(X_1,t) = \int dN_2^2 \left( N_1^2(1-X_1)N_2^2 \right) J(N_1^2,N_2^2) X_1 N_2^2
\]

(35a)

\( J \) is the Jacobian determinant, whose value is calculated from Eq. (34) to be \( N_1^2 \). Upon substituting the expression (32b) for \( r \) into Eq. (34a), and performing the integral, we obtain

\[
R(X_1,t) = \int \left( X_1 \exp(-16\lambda^2) X_1 \right) \frac{1}{(1-X_1)}
\]

(35b)

Eq. (35b) describes a solution whose initial distribution \( \delta(X_1 - X_{10}) \) breaks up on a time scale \( \lambda^{-1} \) 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

\[
\partial R(X_1,t)/\partial t = 2\lambda \Sigma_k \left( \partial \alpha_k \psi_k \right)^2 \psi_k \psi_k \frac{1}{(1-X_1)}
\]

(37)

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

Eqs. (35b), (36), (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), (36). More recently, Gisin\(^3\) has suggested another theory whose two-state sector is also described by Eqs. (35b), (36). We have proved\(^1\) 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\(^6\).

7. Density Matrix

The probability density description of the behavior of the ensemble of wavefunctions 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 wavefunction \( \psi_\Omega(x,t) \), it is necessary to normalize it first, obtaining \( \psi_\Omega(x,t)/N_\Omega(t) \). The probability that this normalized wavefunction is in the physical ensemble is \( \delta N_\Omega^2(0) \). Therefore the density matrix is

\[
\rho(x,t) = \psi_\Omega(x,t) \psi_\Omega^*(x,t) \]

In Appendix A, Eq. (A.15), we give the generalization of Eq. (36) to any number of packets.
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 \( \psi \) we have already found the expression (16) for \( \langle \psi(x,t) \psi^*(x',t) \rangle \). Taking the expectation value of Eq. (16) we obtain:

\[
\mathcal{D}(x,x') \frac{\partial}{\partial t} = \{-i[H(x) - H(x')] + \lambda[\Phi(x-x') - 1]\} \mathcal{D}(x,x')
\]  

(39)

Eq. (39) clearly shows how the usual Hamiltonian evolution occurs for density matrix elements in the position representation taken between nearby points \( |x-x'| < \alpha^{-1/2}, \Phi = 1 \), but is modified by the exponential decay of off-diagonal matrix elements between distant points \( |x-x'| > \alpha^{-1/2}, \Phi = 0 \).

Of course, Eq. (39) can also be obtained from the probability density (see Appendix A, Eqs. (A.12), (A.13)).

Eq. (39) is identical to the equation of evolution of the density matrix proposed by GRW. However, we wish to emphasize that the behavior of the ensemble of wavefunctions 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 \( R(X_1,X_{10},x) \) is given by Eq. (35b). The Poisson process probability density \( R_{GRW}(X_1,X_{10},x) \) is obtained in Appendix B, Eq. (B.13), as

\[
R_{GRW}(X_1,X_{10},x) = \delta(x_1 - x_{10})\exp(-\lambda t) + R(X_1,X_{10}\alpha(x_2-x_1)^2)[1 - \exp(-\lambda t)]
\]

- \( \delta(x_1 - x_{10})\exp(-\lambda t) \)
- \( \delta(x_1 - x_{10})\exp(-\lambda t) \)

Both probability densities \( R \) and \( R_{GRW} \) have initial value \( \delta(x_1 - x_{10}) \) and final value \( X_{10}\delta(1 - x_1) + (1 - X_{10})\delta(x_1) \). However, for \( t < \alpha^{-1} \), \( R \) describes an ensemble in which a negligible amount of wavefunction reduction has occurred i.e. there is a negligible probability that \( X_1 = 1 \) 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 wavefunctions (last term in Eq. (40b) \( \sim \lambda t \) ) for small \( t \).

8. Distinguishable Particles

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

\[
\frac{\partial \psi(x_1,x_2,...,x_n)}{\partial t} = -i[H(x) - H(x')] + \sum_{k=1}^{n}[\delta(x_k - x_{k'}) - n]\psi(x_1,x_2,...,x_n)
\]

(41)

First we look at the density matrix evolution equation. \( \psi \) can be calculated from Eq. (41) just as Eq. (16) was obtained from Eq. (4) in section 4:

\[
\frac{\partial \mathcal{D}(x_1,x_2,...,x_n,x',x'_n)}{\partial t} = -i[H(x) - H(x')]\mathcal{D}(x_1,x_2,...,x_n,x',x'_n) + \sum_{k=1}^{n}[\delta(x_k - x_{k'}) - n]\mathcal{D}(x_1,x_2,...,x_n,x',x'_n)
\]

(43)

Remembering the result of section 7, that \( \mathcal{D}(x_1,...,x_n;x_1',...x_n') = <\psi(x_1,...,x_n)|\psi(x_1',...x_n')> \), by taking the expectation value of Eq. (43) we obtain the density matrix evolution equation

\[
\frac{\partial \mathcal{D}(x_1,x_2,...,x_n)}{\partial t} = -i[H(x) - H(x')]\mathcal{D}(x_1,x_2,...,x_n) + \sum_{k=1}^{n}[\delta(x_k - x_{k'}) - n]\mathcal{D}(x_1,x_2,...,x_n)
\]

(44)

Eq. (44) is precisely the density matrix evolution equation proposed by GRW.

Following them, consider an initial wavefunction describing a macroscopic body whose center of mass \( \mathbf{R} = n^{-1}\sum x_k \) (we take all masses equal for simplicity) has an arbitrary probability distribution, but whose relative coordinates \( r_k = x_k - \mathbf{R} \) are each well localized within a spherical volume of radius \( \epsilon \alpha^{-1/2} \). Since \( \Phi(x_k - x_{k'}) = \Phi(r_k - r_{k'} - \mathbf{R} - \mathbf{R}') \) and when \( |\mathbf{R} - \mathbf{R}'| \) is comparable in size to \( |r_k - r_{k'}| \) we have \( \Phi = 1 \), we may replace each \( \Phi(x_k - x_{k'}) \) in Eq. (44) with \( \Phi(|r_k - r_{k'}|) \) to a good approximation. Then Eq. (44) (with \( H = 0 \)) becomes

\[
\frac{\partial \mathcal{D}(r,r',r'',r'''|\mathbf{R},\mathbf{R}',\mathbf{R}'',\mathbf{R}''')}{\partial t} = -i[H(r) - H(r')]\mathcal{D}(r,r',r'',r'''|\mathbf{R},\mathbf{R}',\mathbf{R}'',\mathbf{R}''')
\]

(45)

Eq. (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 \( r_k \) or over the center of mass coordinate \( R \): the reduced density matrix \( D(R,R';t) \) satisfies Eq. (45), while the other reduced density matrix satisfies \( \partial D(r,r';t)/\partial t = 0 \).

Of course these results for the density matrix have their counterparts in the probability density description of the ensemble of wavefunctions. Let us examine the behavior of the squared amplitudes of the unrenormalized wavefunction 

\[ Y(x_1,...,x_n|t) = |\psi(x_1,...,x_n|t)|^2 \]

just as was done for a single particle moving in one dimension in section 4. Setting \( x_k = x_k' \) (and \( H = 0 \)) in Eq. (43) results in the basic equation of motion for \( Y \)

\[ dY(x_1,...,x_n) = 2\Sigma_k dw_k(x_k,t)Y(x_1,...,x_n|t) \]  

(46)

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

\[ < dw_k^* (x_k,t) dw_l^* (x_l',t) > = \delta_{kl} \lambda \delta (|R - R'|) dt \]  

(47)

Therefore, from Eqs. (46), (47), the diffusion coefficient for \( Y(r_1,...,r_{n-1},R;t) = Y(x_1,...,x_n|t) \) is

\[ < dY(r,R) dY(r',R') > = 4\lambda n \delta (|R - R'|) Y(r,R;t) Y(r',R';t) \]  

(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's look at the behavior of the center of mass. By integrating Eq. (48) over the relative coordinates \( r_k \), \( k = 1,...,n-1 \) we obtain

\[ < dY_{\text{cm}}(R,R';t) > = 4\lambda n \delta (|R - R'|) Y_{\text{cm}}(R,t) Y_{\text{cm}}(R';t) dt \]  

(49)

where \( Y_{\text{cm}}(R,t) = \int dr_{n-1} Y(r,R,t) \) is the squared amplitude for the center of mass alone.

Eq. (49) is identical in form to the diffusion coefficient for a single particle given in Eq. (22), except that \( \lambda \) in Eq. (22) is replaced by \( \lambda n \), and one spatial dimension is replaced by three. Therefore the whole single particle behavior discussed in sections 4, 5, 6 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 \( Y(r,R;t)/Y(r',R';t) \) as was done in section 4. The result has the same form as Eqs. (28), (29):

\[ < dY/Y^* > = 4\lambda n (1 - \Phi(|R - R'|)) dt \]  

(50a)

\[ < d(Y/Y^*)^2 > = 6\lambda n (1 - \Phi(|R - R'|)) dt \]  

(50b)

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

Extending the GRW Poisson process to many indistinguishable particles has proved to be not an easy task. 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

\[ d < x_1,...,x_n | \psi,> = [ -iH + \Sigma_k dw_k(x_k,t) - (1/2) \Sigma_k dw_k(x_k,t)^2 ] < x_1,...,x_n | \psi,> \]  

(51a)

\[ = [ -iH + \Sigma_k dw_k(x_k,t) - (1/2) \Sigma_k dw_k(x_k,t) ] < x_1,...,x_n | \psi,> \]  

(51b)

The statistical properties of \( dw_k(x,k) \) are described by Eq. (5) as usual, (with \( \Phi(x,x') \) replaced by \( \Phi(x,x') \)). Unlike Eq. (41) for distinguishable particles, only one Brownian function \( w(x_k,t) \) is needed.

The calculation of the evolution equation for the density matrix proceeds as in the last two sections. One calculates \( d[Y(x,t) \psi^*(x',t)] \) using Eq. (51) and the Ito rule (15) and, upon taking the expectation value, one obtains

\[ \partial D(x_1,...,x_n;x_1',...,x_n'|t)/\partial t = [ -i[H(x) - H(x')] + \lambda \Sigma_k ( \Phi'(x_k - x_k') (1/2) \Phi(x_k - x_k') - (1/2) \Phi(x_k' - x_k') ) D(x,x') ] \]  

(52)

After having tantalizingly displayed the basic equations (51), (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.

9. Concluding Remarks

The equations of motion presented here (Eqs. (4) or (41) or (51)) describe a
nonunitary but linear evolution of the statevector. Subject to a particular Brownian function $w_Q(x,t)$, a sum of two statevectors at time 0 becomes, at time $t$, a statevector which is the sum of the two evolved statevectors. However, the theory is nonlinear in its rule of weighting each statevector's importance in the ensemble by the squared norm of the statevector. 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.3

Along with the preferred basis and trigger problems mentioned in section 1, 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 statevector 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 have described a conflict of relativity with quantum theory - plus - instantaneous statevector 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 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 statevector reduction occurs. Then in other Lorentz frames there will be intervals of time over which the sum of squared amplitudes is either 1/2 or 3/2. That is, the norm of the statevector 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 statevector 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 statevector reduction.20,21,22,23 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 has shown how to obtain the coupled Schrödinger equation and Einstein equation from an action principle, by varying the statevector 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), 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 statevector, this embarrassing conflict of semiclassical gravity with experiment would disappear.
Acknowledgments

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 a faculty fellowship, 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 wavefunction \( \psi(x,0) \) into a physical ensemble of normalized wavefunctions.

First consider the behavior of squared amplitudes at \( M \) points uniformly spaced along the \( x \)-axis, with spacing \( \Delta \). We define

\[
Z_m(t) = \frac{1}{\Gamma(x_m,t)^2} \int \psi(x_n,t)^2 \psi(x_m,t) \, dx_n(t) \frac{\partial}{\partial Y(t)} \left( Y_n(t) Y_m(t) \right)
\]

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

\[
\sum_m \Delta Z_m = 1
\]

(\( A.2 \))

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

\[
\frac{\partial p(Y,t)}{\partial t} = 2 \lambda \sum_{m,n} \Phi(x_n - x_m) \left( \frac{\partial}{\partial Y_n} \right) \left( \frac{\partial}{\partial Y_m} \right) Y_n Y_m p
\]

(A.3)

for the probability density \( p(Y,t) \) of the unnormalized squared amplitudes \( Y_n \) 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.

Eq. (A.2) 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 \( \Delta^{-1} \) as an independent variable along with \( Y \), and \( Y \) as an independent variable along with \( Z \). We replace \( p \) by

\[
p(Y,t) \to p(Y,\Delta^{-1},t) = p(Y,t) \delta(\Delta^{-1} - \Delta_0^{-1})
\]

(A.4)

which still satisfies Eq. (A.3). The probability density of \( Z,Y \)

\[
W(Z,Y,t) dZ dY = \delta(\Delta^{-1} - \Delta_0^{-1}) dY d\Delta^{-1}
\]

(A.5)

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

\[
\int W dZ dY = \Delta_0 \sum_n Y(t)> = \Delta_0 \sum_n Y(0)
\]

(A.6)

the last step following from the Martingale nature of each \( Y_n \). The \( \delta(\Delta^{-1} - \Delta_0^{-1}) \) factor in Eq. (A.5) approaches the squared wavefunction norm in the continuum limit, and therefore is the right factor to correct the raw ensemble to the physical ensemble in that limit.

Actually, we are most interested in the probability distribution of the \( Z_m \)'s.
regardless of \( Y \):
\[
W(Z,t) = \int dY W(Z,Y;t) \delta(\Sigma_n Z_n - \Delta_0^{-1}) J(Z_0 Y, \Delta_0^{-1} | Z,Y) \quad (A.7)
\]

obtained by integrating Eq. (A.5) over \( Y \), and by using Eqs. (A.1), (A.2) to write \( \Delta^{-1} \) in terms of \( Z \), \( Y \). In Eq. (A.7), \( J(Y, \Delta^{-1} | Z,Y) \) is the Jacobian of the transformation from \((Y, \Delta^{-1})\) to \((Z,Y)\), and is readily calculated to be

\[
J(Y, \Delta^{-1} | Z,Y) = (Y/\Sigma_n Z_n)^M \quad (A.8)
\]

Now, to obtain the Fokker-Planck equation for \( W \), take the derivative of Eq. (A.7) with respect to \( t \). Substitute Eq. (A.3) for \( \partial \rho/\partial t \) into the right hand side of the resulting equation. Use

\[
\partial \rho_n = (S/\gamma) \partial \rho_n - (1/\gamma) \Sigma_m Z_m \partial \rho_m + \partial \rho Y
\]

\((S = \Sigma_m Z_m)\) which follows from Eq. (A.1), to replace derivatives with respect to \( \rho_n \)'s by derivatives with respect to \( Z_m \)'s. After integration over \( Y \), and after some manipulation we obtain

\[
\partial W(Z,t)/\partial t = 2 \Sigma_m Z_m \Phi(x_n, x_m) \alpha
\]

\[\times \left[ (S/\gamma) \partial \rho_n - (1/\gamma) \Sigma_m Z_m \partial \rho_m + \partial \rho Y \right] \quad (A.9)
\]

After further manipulation, Eq. (A.10) can be put in the forms

\[
\partial W(Z,t)/\partial t = 2 \Sigma_m Z_m \partial \rho_n (\partial \rho_n)_{nm} (\partial \rho_n)_{mn}
\]

\[\times (S/\gamma) \partial \rho_n - (1/\gamma) \Sigma_m Z_m \partial \rho_m + \partial \rho Y \quad (A.11a)
\]

\[
\mu_{nm}(Z) = Z_n Z_m \Phi_{nm} + S^2 \Sigma_k \Phi_{nk} Z_k \cdot S^2 \Sigma_k \Phi_{nk} Z_k \quad (A.11b)
\]

\[
\Phi_{nm} = \Phi(x_n, x_m) \quad (A.11c)
\]

Eq. (A.11) describes a gambler's ruin game among the \( Z_m \)'s. Since any function of \( S = \Sigma_k Z_k \) commutes with the differential operators \( \partial \rho_n (\partial \rho_m) \) in Eq. (A.11b), if \( \rho \) is initially proportional to \( \delta(\Delta S - 1) \) it remains so. Thus the amount of "money" in the "game" is preserved.

Likewise, it follows from Eq. (A.11) that \( \partial <Z_n>/\partial t = 0 \), so each \( Z_n \) is a Martingale, and the game is "fair." The complicated dependence of the diffusion constants on the \( Z_m \)'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 \( |x_i - x_j| \ll a^{-1/2} \). 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:

\[
\partial W(Z,t)/\partial t = 2 \Sigma_m Z_m \Phi(x_n, x_m) \alpha
\]

\[\times \left( [dx_1 (\Phi(x_n,x_1) + \Phi(x_1,x_2)) Z(x_1) - \Phi(x_n,x_2)] [dx_1 dx_2 \Phi(x_1,x_2) Z(x_1) Z(x_2)] J(W) \right) \quad (A.11d)
\]

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

\[
D(X,t) = \int dZ(Z,t) |\psi(x,t)|^2 \quad (A.12)
\]

\[\phi(x,t) = (2\pi)^{-1} \delta(x,t) \int dZ(Z,t) |Z(x,t)|^2 \quad (A.13)
\]

which was obtained much more easily in section 7, Eq. (39).

It is easy to apply Eq. (11b) to the situation of \( K \) wavepackets, each of width \( << a^{-1/2} \) separated by distances \( >> a^{-1/2} \). We define the squared norm of the \( k \)th packet as

\[
X_k(t) = k \Sigma_m Z_m \to k \Sigma_m Z_m \quad (A.14)
\]

where the subscript \( k \) restricts the sum or integral to the \( x \)-interval of support of the \( k \)th packet. The arrow refers to the continuum limit where \( Z_m \to Z(x_m) \) is the square of the normalized wavefunction at \( x_m \).

To find the probability density \( R \) of the packet squared norms

\[
R(X_k) = \int dZ(Z,t) |\psi(x,t)|^2 \delta(X_k - X_k) \quad (A.15)
\]

we take the derivative of Eq. (15) with respect to \( t \), and substitute Eq. (11b) for \( \partial W/Z \). The diffusion coefficients (A.11c) in this situation are

\[
\mu_{nm} = Z_n Z_m \cdot (\delta_{nm} Z_n - \delta_{mn} Z_m) \quad (A.16)
\]

where we have used \( \Phi_{nm} = 1 \) or 0 depending upon whether \( x_n \) and \( x_m \) are in the...
same packet or not. In Eq. (A.16), 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 $X_n$'s by means of the delta functions in Eq. (A.15), the following result is obtained:

$$
\frac{dR(X_1,t)}{dt} = 2\lambda \sum_k \delta(X_1 - X_k) \delta(X_1 - X_k - \Sigma X_k^2) R(X_1) \tag{A.17}
$$

Eq. (A.17) describes a gambler's ruin game among the squared packet norms $X_k$.

Its significance is discussed in section 6.

---

**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

$$
X_k(t) = \int_0^t dx |\psi(x,t)|^2 \quad (k=1 \text{ or } 2) \tag{B.1}
$$

is the squared norm associated with packet $k$ belonging to the normalized wavefunction $\psi$ (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 wavefunction will be hit. If $\psi$ is hit with a center at $z$, the wavefunction immediately after the hit is

$$
(\alpha/\pi)^{1/4} \exp(-(\alpha/2)(x-z)^2) \psi(x,t)/\sqrt{N(z,t)} \tag{B.2}
$$

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

$$
N^2(z,t) = (\alpha/\pi)^{1/2} dx |\exp(-\alpha(x-x)^2)| |\psi(x,t)|^2 \tag{B.3a}
$$

$$
= (\alpha/\pi)^{1/2} (X_1|\exp(-\alpha(x_1-x)^2)| + X_2|\exp(-\alpha(x_2-x)^2)|) \tag{B.3b}
$$

In obtaining Eq. (B.3b) from (B.3a) 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. (B.2) and inserting it into (B.1), we obtain the result that a packet characterized by $X_1$ before a hit becomes characterized by

$$
X = a_1 X' / [a_1 X + a_2(1-X)] \tag{B.4a}
$$

$a_k = \exp(-\alpha(x_k-x^2)^2)$ after the hit. The inversion of Eq. (B.4a) is

$$
X' = a_2 X / [a_2 X + a_1(1-X)] \tag{B.4b}
$$

The probability $R(X_1+dt)dX$ 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 (B.3b),

$$
\lambda dt (\alpha/\pi)^{1/2} \int dz [a_1 X + a_2(1-X)] R(X_1) dX' \tag{B.5a}
$$

where it is understood that $X'$ in Eq. (B.5a) is the expression (B.4b), which is a function of $z$ and $X$. The probability that there is a transition out of this range is

$$
\lambda dt (\alpha/\pi)^{1/2} \int dz [a_1 X + a_2(1-X)] = \lambda dt \tag{B.5b}
$$
Therefore the equation of evolution for $R(X;\tau)$ is

$$\partial R(\tau) = \lambda(\partial \alpha / \partial \tau)^{1/2} \int_0^\tau dz \frac{\partial z}{\partial X} R(\tau;\tau) \frac{\partial X}{\partial z} - \lambda R(\tau) \tag{B.6}$$

In order to solve Eq. (B.6), it is useful to change the variable of integration from $z$ to $X$. We note from Eq. (B.4b) that, as $z$ ranges from $-\infty$ to $\infty$, $X$ ranges from $0$ to $1$ (with the choice $x^2 \leq x_1$). By use of $(\partial z / \partial X)^{-1} \frac{\partial X}{\partial z} = -\frac{\partial (\partial z / \partial X)}{\partial X}$, Eq. (B.6) becomes

$$\partial R(X;\tau) = \lambda(\partial \alpha / \partial \tau)^{1/2} \int_0^\tau dX \left[ a_1(1-X^2) + a_2(1-X^2) \right] R(X;\tau)$$

$$- \lambda R(X;\tau) \tag{B.7a}$$

By solving Eq. (B.4) for $z$ in terms of $X,X'$ we obtain

$$a_1(z(X,X')) = [X(1-X)/X'(1-X)]^{1/2} \tag{B.7b}$$

Remarkably, it turns out that $Xa_1$ multiplied by the bracket outside the integral in Eq. (B.7b) is $G(X,X';\tau)$, where $G(X,X';\tau)$ is the solution Eq. (35) of Eq. (36):

$$\partial G(X,X';\tau) = \frac{\partial \alpha}{\partial X} \left[ X^2(1-X)^2 \frac{\partial G(X,X';\tau)}{\partial X} \right]$$

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

$$\partial R(X;\tau) = \lambda \int_0^\tau dX' G(X,X';\tau) [X^2(1-X)^2] R(X';\tau) - \lambda R(X;\tau) \tag{B.10}$$

We can now proceed to solve Eq. (B.10). The explicit expression for $G(X,X';\tau)$ appears in Eq. (35) (with $4X \rightarrow 1, T \rightarrow T$, $X_1 \rightarrow X$, $X_10 \rightarrow X'$). Since $\alpha(x_2-x_1)^2 \geq 1$, we only need to know $G$ for large $T = \alpha(x_2-x_1)^2$:

$$G(X,X';\tau) = X(1-X) + (1-X)X' \tag{B.11}$$

The $\delta$-functions in Eq. (B.11) are actually narrow peaks of height $\sim e^{T}$ and of area close to $1$, centered at $X = [1+\alpha(1-X)^{-1}]^{-1}$. Putting the expression (B.11) for $G$ into Eq. (B.10) yields

$$\partial R(X;\tau) = \lambda \alpha(X' R(\tau;\tau) - R(X)) \tag{B.12}$$

The solution of Eq. (B.12), subject to the initial condition $R(X;0) = \delta(X-X_0)$ is

$$R(X) = \delta(X-X_0) \exp(\lambda \tau) + [X_0 \delta(1-X) + (1-X) \delta(X)] [1 - \exp(-\lambda \tau)]. \tag{B.13}$$

The significance of this solution is discussed in section 7.

---

Appendix C

In this Appendix we explain how the probability weighting of the physical ensemble is consistent with an independent Markovian evolution for each statevector in the ensemble.

Consider an ensemble of statevectors, each of which evolves for $0 \leq \tau < T$, according to the Schrödinger equation, by means of one of a set of possible non-Hermitian Hamiltonians $H_j$. Suppose that a particular statevector "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 $j^{th}$ Hamiltonian is chosen might depend upon the time interval, upon the history and present state of the statevector making the choice, and indeed upon the histories and present states of all the other statevectors 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 $T \to 0$, 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 statevector 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 $T \to 0$, it is truly Markovian.)

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

The probability rule can be generalized without losing these three properties. If the rule depends upon the statevector making the choice as well as upon the index $j$, that
statevector still evolves independently. If in addition the rule is independent of the time
interval, and depends only upon the statevector 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 statevector which has evolved to the beginning of a particular
time interval. Generally that statevector's norm will not equal 1. So, normalize it.
Now, consider the evolution \( |\psi(0)\rangle \rightarrow |\psi(r)\rangle \) of that normalized statevector over the
time interval \( T \) by means of the Hamiltonian \( H_j \), i.e. \( i\hbar\psi(t)/dt = H_j \psi(t) \). The statevector's
squared norm at time \( T \) is \( N_j^2 = \langle \psi_j(T)|\psi_j(T)\rangle \). Our rule is \( P_j = N_j^2/\Omega \).

Note that \( N_j^2 \), and therefore \( P_j \), depends only upon the statevector 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 statevector
\( |\psi\rangle \),
\[
\Sigma P_j = \int_j \sum N_j^2 = \int N_j^2 = 1 \quad \text{for any } |\psi\rangle \tag{C.1a}
\]
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
\[
\int_j \sum V_j(T)^* V_j(T) = 1 \tag{C.1b}
\]
and we will suppose this is the case.

Finally, we can obtain an interesting result concerning the norms of an ensemble
of statevectors at time \( t=NT \) which evolved from a single statevector \( |\psi,0\rangle \).

Consider one of the evolved statevectors \( |\psi_{\Omega,NT}\rangle = |\psi_{\Omega,n}(n)\rangle_{[1]} \) in the ensemble
which utilized the \( j(n)^{th} \) Hamiltonian at the \( n^{th} \) time interval. The probability that it is in
the ensemble is
\[
P(|\psi_{\Omega,NT}\rangle) = \prod_j \int_{n=1}^{N} P_j(n) = \int_j \prod_{n=1}^{N} N_j(n)^2 [j(n-1)...[1]];|\psi,0> \tag{C.2}
\]
(We have indicated in Eq. (C.2) that the squared norm \( N_j(n)^2 \) acquired during the \( n^{th} \)
interval depends not only on \( j(n) \), but also on the statevector at the beginning of the \( n^{th} \)
interval, and that statevector in turn depends upon the initial statevector and its
subsequent history, up to that interval.)

On the other hand, the norm of this statevector is
\[
\langle \psi_{\Omega,NT}\rangle_{[1]} = \langle \psi_{\Omega,NT}|\psi_{\Omega,NT}\rangle = \langle \psi_{\Omega,NT}|V_j(T)^*.. V_j(N(T)^* V_j(N(T).. V_j(1(T)|\psi,0\rangle
\]
\[
= \prod_{n=1}^{N} N_j(n)^2 [j(n-1)...[1]];|\psi,0> \tag{C.3}
\]
Combining Eqs. (C.2) and (C.3) we obtain the result
\[
P(|\psi_{\Omega,NT}\rangle) = \int_j \langle \psi_{\Omega,NT}|\psi_{\Omega,NT}\rangle \tag{C.4}
\]
Thus the probability that the statevector \( |\psi_{\Omega,NT}\rangle \) belongs to the ensemble at time \( NT \)
is proportional to the norm of that statevector. The proportionality factor \( \Omega \) 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 \( T \rightarrow 0 \) 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. (C.4) becomes
the probability rule \( P_{\Omega} = d\Omega \Omega^2 \) of the physical ensemble.
References and Remarks

8. In the GRW theory and in the theory presented here the statevector is never completely reduced. There is always a small but nonvanishing piece of "what might have been" included in the statevector. We do not regard this as satisfactory. If the reduced statevector 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 references 3 and 12, pointing out that the theory described in reference 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 \( \int dx(x,t) \int dx'(x) dw(x,t) > 0 \) for arbitrary \( t \neq 0 \), \( \Phi \) must be positive definite, i.e., \( \int \text{d}x(x) \Phi(x-x') \Phi(x') > 0 \). Any such function will do which also has \( \Phi(0) = 1, \Phi(\pm) = 0 \), and a characteristic length governing the transition of \( \Phi \) 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 \( \phi(x,t) = (2\pi)^{-1/2} \exp[-i(x-x)^2/(2\sigma^2)] \), we calculate \( \langle \chi \rangle = 0 \), \( \langle \chi \chi \rangle > 0 \) using Eqs. (4), (5) and Eqs. (15), (26). Since \( \phi(x,t) \) has vanishing drift and diffusion, it does not change with time.
13. It has, however, been recently accomplished (G. C. Ghirardi, P. Pearle and A. Rimini, private communication).
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 \( \psi \). For example, Eq. (4) is \( \text{d} \psi(x,t) / \text{d} t = -i[H \psi(x,t) / \text{d} t + \text{d} \omega(x,t) / \text{d} t - \lambda \psi(x,t)] \) in Stratonovich form (i.e. the 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 \( \text{d} \omega(x,t) / \text{d} t \) 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).
23. D. W. Sciama (private communication).