

Reduction of the state vector by a nonlinear Schrödinger equation

Philip Pearle

Hamilton College, Clinton, New York 13323

(Received 14 April 1975; revised manuscript received 9 September 1975)

It is hypothesized that the state vector describes the physical state of a single system in nature. Then it is necessary that the state vector of a macroscopic apparatus not assume the form of a superposition of macroscopically distinguishable state vectors. To prevent this, it is suggested that a nonlinear term be added to the Schrödinger equation, which rapidly drives the amplitude of one or another of the state vectors in such a superposition to one, and the rest to zero. It is proposed that it is the phase angles of the amplitudes immediately after a measurement which determine which amplitude is driven to one. A diffusion equation is arrived at to describe the reduction of an ensemble of state vectors corresponding to an ensemble of macroscopically identically prepared experiments. Then a nonlinear term to add to the Schrödinger equation is presented, and it is shown that this leads to the diffusion equation in a weak-coupling approximation.

I. INTRODUCTION

What is it in Nature that is in one-to-one correspondence with the state vector of quantum theory? Most physicists answer: "The state vector corresponds to an ensemble of identical systems in Nature." Bohr¹ felt that quantum theory, with this interpretation, was a complete and satisfying theory. However, Einstein^{2,3} argued that, since single systems exist in Nature and quantum theory does not describe them, quantum theory is an incomplete description of Nature and should be modified.

In the present form of quantum theory, the statement "The physical state of a single system in nature is in one-to-one correspondence with the state vector of the theory" is not acceptable. In this paper we propose an altered form of quantum theory for which this statement *is* acceptable. Then this theory can be regarded as directly describing reality, and need not be thought of as describing just the statistical behavior of reality.

The conflict between the above statement and quantum theory was made evident by Schrödinger⁴ with his "cat paradox." As is well known,⁵ Schrödinger pointed out that if one accepts the truth of the above statement, then it is possible to set up an experimental arrangement whereby the physical state of a single cat in a closed box is in one-to-one correspondence with the state vector

$$\frac{1}{\sqrt{2}} |\text{cat alive}\rangle + \frac{1}{\sqrt{2}} |\text{cat dead}\rangle. \quad (1.1)$$

This conflicts with our common sense notion that a cat (whether observed or not) is either alive or dead, and so its physical state should be in correspondence with either of the state vectors

$$|\text{cat alive}\rangle \text{ or } |\text{cat dead}\rangle. \quad (1.2)$$

In other words, the state vector (1.1) predicted by

quantum theory corresponds to no physical state of the cat, and this contradicts the statement. Therefore, either the statement is false, or our common sense is at fault, or quantum theory is at fault. It is usual to reject the statement. However, in this paper we will accept the statement, and examine the possibility that quantum theory is at fault.

It is the *linearity* of the Schrödinger equation which produces the state vector (1.1).⁶ If quantum theory is modified so that the Schrödinger equation becomes nonlinear, it could be possible to produce a final-state vector of the cat which is one of those in (1.2). However, when the experiment is repeated many times, half the time the cat comes out alive and half the time it comes out dead. In order to agree with the outcome of this repeated experiment, the nonlinear Schrödinger equation must produce different solutions each time it is solved: Half the time it must predict the final-state vector $|\text{cat alive}\rangle$, and half the time it must predict the final-state vector $|\text{cat dead}\rangle$. But since the final-state vector predicted by the nonlinear Schrödinger equation must be uniquely determined by the initial values of the variables in the equation, we must answer the question "What are the variables in the nonlinear Schrödinger equation whose differing initial values lead to differing final-state vectors?"

Recently, Bohm and Bub⁷ have introduced a Schrödinger equation modified by nonlinear terms which depend on certain "hidden variables" first introduced by Wiener and Siegel.⁸ It is the initial values of these hidden variables which determine what the final-state vector will be. Such a theory has been called a "hidden-variable theory of the first kind" by Belinfante.⁹

In the modification of the Schrödinger equation suggested here, no new variables are to be introduced into quantum theory. The variables will

be the amplitudes $x_n^{1/2}$ and phases θ_n of the coefficients of the expansion of the state vector $|\psi(t)\rangle$ in a certain basis $|\varphi_n(t)\rangle$ (to be prescribed later):

$$|\psi(t)\rangle = \sum_n x_n^{1/2} \exp[i\theta_n(t)] |\varphi_n(t)\rangle. \quad (1.3)$$

It will be the initial values of the θ_n which determine the final-state vector. Loosely speaking, the phase angles θ_n are the hidden variables of this theory.

The theory is to work in the following way. Consider an experiment wherein a microscopic system interacts with a macroscopic system (apparatus). Before the interaction takes place, the combined system (hereafter simply called "the system") is to be described by a state vector, say $|\chi(t)\rangle$. However, the correct state vector to use in describing the system is never precisely known because the physical state of the system can never be precisely defined experimentally. As a result, one can construct an ensemble of state vectors $|\chi_k(t)\rangle$, each one of which is consistent with all known experimental facts about the system. One does not know which state vector to choose to describe the system, so one may regard them *a priori* as equally likely.

Choose one of these initial-state vectors, and follow its time evolution. Before the interaction, $|\psi(t)\rangle = |\chi_k(t)\rangle$. Immediately after the interaction, $|\psi(t)\rangle$ will have the form (1.3), where each $|\varphi_n(t)\rangle$ corresponds to a possible outcome of the experiment. The Schrödinger equation without the nonlinear term would maintain (1.3) as a superposition of macroscopically different states, with the values of x_n essentially unchanged, after the interaction has ceased. However, the nonlinear Schrödinger equation (whose effects should be negligibly different from the effects of the linear Schrödinger equation during the interaction) should rapidly drive all x_n 's to zero, except one x_n which is driven to unity, after the interaction has ceased. If this occurs, we will say that the nonlinear Schrödinger equation possesses "property 1."

Thus, with a nonlinear Schrödinger equation possessing property 1, each initial-state vector $|\chi_k(t)\rangle$ determines a final-state vector $|\varphi_n(t)\rangle$ which corresponds to a particular macroscopic outcome of the experiment. The *indeterminism* that is observed in Nature is regarded as due to our inability to select the precisely correct state vector corresponding to the initial physical state of the system. Another initial-state vector $|\chi_i(t)\rangle$, just slightly different from $|\chi_k(t)\rangle$, can result in a completely different final-state vector $|\varphi_m(t)\rangle$. In order that we obtain agreement with the statistical predictions of quantum theory, the nonlinear Schrö-

dinger equation must have a second property which determines how frequently each final-state vector $|\varphi_n(t)\rangle$ occurs.

We proceed to define this second property by first grouping the initial-state vectors into equivalence classes as follows. Starting with any initial-state vector, compute $|\psi(0)\rangle$, where $t=0$ is a time immediately after the interaction has ceased, but before the reduction (the effect of the nonlinear term) has begun. $|\psi(0)\rangle$ will have the form (1.3) All initial-state vectors for which $|\psi(0)\rangle$ has identical $x_1(0), \dots, x_n(0), \dots$ [but different $\theta_1(0), \dots, \theta_n(0), \dots$] are in the same equivalence class. We will assume that, in each equivalence class, every $\theta_n(0)$ is randomly distributed. It is the values of these $\theta_n(0)$, themselves determined by the initial-state vector, which determine the final-state vector.

Agreement with the statistical predictions of quantum theory will be obtained if the nonlinear Schrödinger equation, acting on the initial-state vectors of an equivalence class, produces the final-state vector $|\varphi_n(t)\rangle$ [i.e., drives the n th coefficient $x_n(t)$ to unity] for a fraction $x_n(0)$ of the total number of members of the equivalence class. If this occurs, we will say that the nonlinear Schrödinger equation possesses "property 2."

To summarize: The reason we get different results each time we repeat an experiment is attributed to the fact that each time we start out with a different physical state of the system, and therefore a different initial-state vector, obtain different phase angles $\theta_n(0)$ after each interaction, and so each time the nonlinear Schrödinger equation drives a different x_n to unity.

Our approach to the construction of such a theory proceeds in two steps. In the first step, in Secs. II and III, we construct a "diffusion equation" or "master equation." This equation is to statistically describe the reduction behavior of the solutions of a nonlinear Schrödinger equation possessing properties 1 and 2. Then in Sec. IV we suggest a nonlinear term to add to the Schrödinger equation, and in Secs. V and VI we show that (in suitable approximate schemes of solution) the diffusion equation can be derived from it. The equations introduced are the simplest we have been able to find with the desired properties. We will try to make the choice of these equations plausible, but they are not "inevitable."

II. DIFFUSION EQUATION FOR A TWO-STATE SYSTEM

To illustrate what we have in mind, we will begin by discussing the reduction of a two-state quantum system.¹⁰ We introduce the probability

density $\rho(x_1, x_2, t)$ which gives the probability at time t that the squared amplitudes for the two states will lie in an infinitesimal range about the values x_1, x_2 . If we restrict our attention to the behavior of the state vectors in an equivalence class characterized by certain $x_1(0), x_2(0)$, the initial probability density is

$$\begin{aligned}\rho(x_1, x_2, 0) &= \delta(x_1 - x_1(0))\delta(x_2 - x_2(0)) \\ &= \delta(x_1 - x_1(0))\delta(1 - x_1 - x_2)\end{aligned}\quad (2.1)$$

(of course $x_1 + x_2 = 1$ at any time).

We wish to write down a "diffusion equation" whose solution, subject to the initial condition (2.1), will be the probability density ρ for the

equivalence class. It is to be a linear equation so that the solution for an ensemble of equivalence classes will simply be the sum of solutions for each individual equivalence class. Without more ado, we shall write down the equation we have in mind,

$$\frac{\partial \rho}{\partial t} = \mu \left(\frac{\partial}{\partial x_1} - \frac{\partial}{\partial x_2} \right)^2 (x_1 x_2)^r \rho \quad (2.2)$$

(r is some positive number, μ is a positive coupling constant), present solutions for $r = 2, 1$, and then motivate the choice of this equation.

The solution of (2.2) for $r = 2$, which agrees with (2.1) at $t = 0$ [setting $x_1(0) \equiv x_0$], is

$$\begin{aligned}\rho dx_1 dx_2 &= dx_2 d \ln[x_1(1 - x_0)/(1 - x_1)x_0] \delta(1 - x_1 - x_2) (2\pi 2\mu t)^{-1/2} \\ &\quad \times [x_0 \exp\{-\frac{1}{2}(4\mu t)^{-1}\{-\mu t + \ln[x_1(1 - x_0)/(1 - x_1)x_0]\}^2\} \\ &\quad + (1 - x_0) \exp\{-\frac{1}{2}(4\mu t)^{-1}\{\mu t + \ln[x_1(1 - x_0)/(1 - x_1)x_0]\}^2\}] \\ &= dx_1 dx_2 \delta(1 - x_1 - x_2) [x_0(1 - x_0)x_1^{-3}(1 - x_1)^{-3} 4\pi\mu t]^{1/2} \exp\{-\frac{1}{4}\mu t - (4\mu t)^{-1/2} \ln[x_1(1 - x_0)/(1 - x_1)x_0]\}^2.\end{aligned}\quad (2.3a)$$

$$(2.3b)$$

From (2.3a) we can see that this solution, regarded as a function of x_1 , consists of two peaks, initially superimposed at $x_1 = x_0$, which separate and travel toward $x_1 = 0$ and $x_1 = 1$. The peaks are Gaussian in the variable $\ln[x_1(1 - x_0)/(1 - x_1)x_0]$. The areas under the peaks traveling toward $x_1 = 1$ and $x_1 = 0$ are x_0 and $1 - x_0$, respectively. As $t \rightarrow \infty$,

$$\rho \xrightarrow{t \rightarrow \infty} x_1(0)\delta(x_1 - 1)\delta(x_2) + [1 - x_1(0)]\delta(x_1)\delta(x_2 - 1), \quad (2.4)$$

which exhibits property 1 (embodied in the δ functions) and property 2 (embodied in the numerical coefficients multiplying the δ functions). From (2.3b) it can most clearly be seen that this solution [unlike the solution below in (2.5)] vanishes at $x_1 = 0, 1$ for any finite time.

The solution of (2.2) for $r = 1$, with the same initial condition (2.1), is

$$\rho = \delta(1 - x_1 - x_2) \{U(x_1)U(x_2)f_1(t, x_1, x_0) + \delta(x_1 - 1)[x_0 - f_2(t, 1, x_0)] + \delta(x_1)[1 - x_0 - f_2(t, 0, x_0)]\}, \quad (2.5a)$$

$$\begin{aligned}f_k(t, x_1, x_0) &\equiv [x_0(1 - x_0)/x_1(1 - x_1)]^{1/2} \sum_{n=0}^{\infty} [(2n+3)/(2n+1)^k(2n+2)^k] P_{n+1}^1(2x_1-1)P_{n+1}^1(2x_0-1) \\ &\quad \times \exp[-\lambda(n+1)(n+2)t].\end{aligned}\quad (2.5b)$$

In (2.5), $U(x)$ is the step function ($U = 1$ for $x > 0$; $U = 0$ for $x \leq 0$) and $P_{n+1}^1(t)$ is an associated Legendre function. An important aspect of (2.5) is the presence of the δ functions $\delta(x_1), \delta(x_1 - 1)$ and the step functions $U(x_1), U(x_2)$. When (2.5) is substituted into the right-hand side of (2.2), these δ functions are multiplied by $x_1 x_2 = x_1(1 - x_1)$ and so they vanish [since $x\delta(x)$ and all its derivatives vanish when integrated with respect to infinitely differentiable test functions], while differentiation of the step functions produces δ -function terms obtained when (2.5) is substituted into the left-hand side of (2.2). The process we are considering is somewhat akin to random walk with ab-

sorbing barriers: it is interesting that the pile-up of probability density in the "barriers" at $x_1 = 0$ and $x_1 = 1$ is completely contained in the solution.

As $t \rightarrow \infty$, the solution (2.5) approaches (2.4), so (2.5) also exhibits properties 1 and 2. This of course is no accident; we now turn to the diffusion equation (2.2) to examine the properties it imposes on its solutions.

The dependence of Eq. (2.2) upon the differential operator $(\partial/\partial x_1 - \partial/\partial x_2)$ forces the probability density to flow along lines of constant $x_1 + x_2$. The initial condition (2.1) puts all the probability density on the line $x_1 + x_2 = 1$, so ρ retains the factor $\delta(1 - x_1 - x_2)$ for all t .

To discuss probability conservation, we write $\rho \equiv \bar{\rho}(x_1, t)\delta(1 - x_1 - x_2)$, converting (2.2) into

$$\frac{\partial \bar{\rho}}{\partial t} = \mu \frac{\partial^2}{\partial x_1^2} x_1^r (1 - x_1)^r \bar{\rho}. \tag{2.6}$$

Integration of (2.6) from $x = 0^-$ to 1^+ (\pm refer to the addition or subtraction of an infinitesimal number) yields

$$\frac{\partial}{\partial t} \int \bar{\rho} dx_1 = \mu \left[\frac{\partial}{\partial x_1} (1 - x_1)^r \bar{\rho} \Big|_{x_1=1^+} - \frac{\partial}{\partial x_1} x_1^r \bar{\rho} \Big|_{x_1=0^-} \right]. \tag{2.7}$$

It can be shown that the right-hand side of (2.7) vanishes for $r \geq 2$ because $\bar{\rho}$ is nonsingular at the boundary (in fact, $\bar{\rho}$ and all its derivatives vanish). For $0 < r < 2$, it can be shown that if the limits on the right-hand side of (2.7) were taken from *within* the interval (i.e., $x_1 = 1^-, x_1 = 0^+$), these terms would not vanish: Indeed, they represent the probability flux to the boundary. However, the limits are taken from *outside* the interval, where $\bar{\rho}$ vanishes identically, so probability is conserved. Solutions (2.3) and (2.5) illustrate both kinds of boundary behavior.

To see why $\bar{\rho}$ is non-negative, we rewrite (2.6) as an equation for $\phi \equiv x_1^r (1 - x_1)^r \bar{\rho}$:

$$\frac{\partial \phi}{\partial t} = \mu x_1^r (1 - x_1)^r \frac{\partial^2 \phi}{\partial x_1^2}. \tag{2.8}$$

Because the function multiplying $\partial^2 \phi / \partial x_1^2$ is non-negative on the interval $0-1$, this is an elliptic-parabolic equation. For such an equation there exists a minimum principle,¹¹ stating that ϕ takes on its minimum value on the boundary of the space-time strip $0 \leq x \leq 1, 0 \leq t < \infty$. [Essentially the argument is that a minimum of ϕ within the strip requires $\partial \phi / \partial t = 0, \partial^2 \phi / \partial x_1^2 > 0$, which (2.8) forbids.] It can be shown that ϕ vanishes at the boundary $x_1 = 0, 1$, and as $t \rightarrow \infty$, and by (2.1) is ≥ 0 at $t = 0$. Therefore, $\phi \geq 0$ everywhere and so $\bar{\rho}$ is non-negative.

To show why all solutions possess properties 1 and 2 we multiply (2.6) by x_1 and by $x_1(1 - x_1)$, and after integrations by parts we obtain

$$\frac{\partial}{\partial t} \int_0^1 x_1 \bar{\rho}(x_1, t) dx_1 = 0, \tag{2.9}$$

$$\frac{\partial}{\partial t} \int_0^1 x_1(1 - x_1) \bar{\rho}(x_1, t) dx_1 = -2 \int_0^1 x_1^r (1 - x_1)^r \bar{\rho}(x_1) dx_1. \tag{2.10}$$

The right-hand side of (2.10) will be negative as long as $\bar{\rho}$ has support on the open interval $1 > x_1 > 0$, and will only vanish if

$$\bar{\rho} = a\delta(x_1) + b\delta(1 - x_1). \tag{2.11}$$

Therefore, the integral on the left-hand side of (2.10) is a monotonically decreasing function of time. However, this integral is bounded below by zero, so it approaches a limit as $t \rightarrow \infty$. Then the right-hand side of (2.10) must vanish as $t \rightarrow \infty$, and as this only occurs when $\bar{\rho}$ assumes the form (2.11) we conclude that (2.11) holds in the limit $t \rightarrow \infty$. [Incidentally, the limit of the integral on the left-hand side of (2.10) is therefore also zero as $t \rightarrow \infty$.] This means that the solutions possess property 1.

According to Eq. (2.1), the mean value of x_1 at time $t = 0$ is $x_1(0)$. According to Eq. (2.11), the mean value of x_1 as $t \rightarrow \infty$ approaches b . Since Eq. (2.9) states that the mean value of x_1 does not change, we conclude that $b = x_1(0)$. Similarly, the mean value of $1 - x_1$ does not change, so $a = 1 - x_1(0)$. It then follows from Eq. (2.11) that the asymptotic form of ρ is (2.4), and thus property 2 is obtained.

All these properties of solutions of Eq. (2.2) are unaffected by replacing $(x_1, x_2)^r$ in Eq. (2.2) by any function of x_1, x_2 that is positive for $x_1 + x_2 = 1, 0 < x_1 < 1$, and vanishes on the boundary $x_1 = 0, 1$.

III. DIFFUSION EQUATION

The diffusion equation that generalizes Eq. (2.2) to an N -state quantum system is

$$\frac{\partial \rho(\vec{x}, t)}{\partial t} = \mu \sum_{n < m} \left(\frac{\partial}{\partial x_n} - \frac{\partial}{\partial x_m} \right)^2 \alpha_{nm}^2 x_n^r x_m^r \rho \tag{3.1}$$

(the α_{nm} 's are real constants, and for simplicity in later expressions, we will suppose $\alpha_{nm} = \alpha_{mn}$).

The solutions of Eq. (3.1), subject to the initial condition for an equivalence class

$$\rho(\vec{x}, 0) = \delta(x_1 - x_1(0))\delta(x_2 - x_2(0)) \cdots \delta(x_N - x_N(0)), \tag{3.2}$$

enjoy analogous properties to those of the solutions of Eq. (2.2). We will now state these five properties and sketch the proofs of these statements.

1. ρ is proportional to $\delta(1 - x_1 - \cdots - x_N)$. This follows from the dependence of (3.1) on the differences $(\partial/\partial x_n - \partial/\partial x_m)$, and the vanishing of ρ off the hyperplane $x_1 + \cdots + x_N = 1$ at $t = 0$ due to the initial condition (3.2).

2. Probability is conserved. The integral of (3.1) over a volume slightly larger than the hypercube $0 \leq x_i \leq 1$ results in

$$\frac{\partial}{\partial t} \int_0^1 dx_1 \cdots \int_0^1 dx_N \rho = 0 \tag{3.3}$$

for any solution of (3.1) that vanishes as the boundary of the hypercube is approached from the outside.

3. ρ is non-negative. Multiplying (3.1) by $(x_1 x_2 \cdots x_N)^r$ and writing $\phi \equiv (x_1 x_2 \cdots x_N)^r$, we obtain

$$\frac{\partial \phi}{\partial t} = \mu \sum_{n < m}^N \alpha_{nm}^2 x_n^r x_m^r \left(\frac{\partial}{\partial x_n} - \frac{\partial}{\partial x_m} \right)^2 \phi. \quad (3.4)$$

The right-hand side of (3.4) is an elliptic operator on ϕ on the hyperplane $x_1 + x_2 + \cdots + x_N = 1$ contained within the hypercube $0 < x_i < 1$ because of the positive-definite nature of the characteristic expression

$$\sum_{n < m}^N \alpha_{nm}^2 x_n^r x_m^r (\xi^n - \xi^m)^2 > 0 \quad (3.5)$$

($\vec{\xi}$ is any nonzero vector lying in the hyperplane, i.e., $|\vec{\xi}| \neq 0, \sum_{n=1}^N \xi^n = 0$). Thus by the minimum principle already mentioned,¹¹ ϕ takes on its minimum value on the boundary of the union of the spatial domain with the time domain $0 \leq t < \infty$. For the solutions we are considering, ϕ vanishes on the boundary of the spatial domain and at $t \rightarrow \infty$ and is non-negative at $t = 0$, so $\phi \geq 0$ everywhere.

4. The solution possesses property 1. From (3.1) we have

$$\frac{\partial}{\partial t} \int_0^1 dx_1 \cdots \int_0^1 dx_N x_k \rho = 0, \quad (3.6)$$

$$\begin{aligned} \frac{\partial}{\partial t} \int_0^1 dx_1 \cdots \int_0^1 dx_N x_k (1 - x_k) \rho \\ = -2\mu \int_0^1 dx_1 \cdots \int_0^1 dx_N \left(\sum_{m=k}^N \alpha_{km}^2 x_m^r \right) x_k^r \rho. \end{aligned} \quad (3.7)$$

By the identical argument given in Sec. II, we conclude that the right-hand side of (3.7) must vanish as $t \rightarrow \infty$, which means that ρ is nonvanishing only where the other factors in the integral on the right-hand side of (3.7) vanish, i.e., at $x_k = 0, 1$ ($\sum_{m, m \neq k} \alpha_{km}^2 x_m^r$ vanishes only if each $x_m = 0, m \neq k$, which means $x_k = 1$). Since this holds for all k , ρ is asymptotically the sum of products of δ functions

$$\begin{aligned} \rho \xrightarrow[t \rightarrow \infty]{} c_1 \delta(1 - x_1) \delta(x_2) \cdots \delta(x_N) \\ + \cdots + c_N \delta(x_1) \delta(x_2) \cdots \delta(1 - x_N) \end{aligned} \quad (3.8)$$

(the c_k 's are constants). [Terms such as $\delta(1 - x_k) \delta(1 - x_i)$ cannot occur, since $\rho \sim \delta(1 - x_1 - \cdots - x_N)$.] Thus all the probability density ends up at corners of the hypercube, satisfying property 1.

5. The solution possesses property 2. The expectation value of x_k at $t = 0$ is $x_k(0)$, by (3.2). At $t \rightarrow \infty$ it is c_k , by (3.8). It does not change with time by (3.6). Therefore, $c_k = x_k(0)$.

A few more remarks about Eq. (3.1) conclude this section. These five properties of the solutions of (3.1) are unaffected by replacing each x_k^r by any function of x_k that is positive for $0 < x_k < 1$

and vanishes at $x_k = 0, 1$.

An existence proof does not yet exist, nor have we been able to find exact solutions for $N > 2$.

Equation (3.1) should allow solutions for $N > 2$ which involve step functions and δ functions, as in the solution (2.5). This is discussed further in Appendix A. Also, in Appendix B is a crude estimate of the characteristic time τ for the reduction to take place. It is argued there that $\tau \approx N^{r-1}/\mu$ ($r \geq 1$: N is the number of states). The case $r = 1$, for which τ is independent of N (assuming μ is independent of N) would appear to be of special interest.

IV. NONLINEAR SCHRÖDINGER EQUATION

It has been possible to find a nonlinear Schrödinger equation whose solutions are statistically described by the diffusion equation, in the weak-coupling limit. The strong-coupling limit is as yet imperfectly understood, but might satisfactorily describe the reduction process: This is discussed in Sec. VII. In this section the nonlinear Schrödinger equation will be presented. Then in the next two sections the diffusion equation will be derived in two different ways from the nonlinear Schrödinger equation.

Recalling the discussion in Sec. I, $|\psi(t)\rangle$ is the state vector corresponding to a microscopic system and an apparatus undergoing an experiment. Without the nonlinear term in the Schrödinger equation, we suppose that

$$|\psi(t)\rangle = \exp[-i\hbar^{-1}(H_0 + H_I)t] |\psi(0)\rangle, \quad (4.1)$$

where $H_0 + H_I$ is the usual quantum-mechanical Hamiltonian for the combined system. It is assumed that the effect of H_I is negligible except for a brief interval of time prior to $t = 0$ during which the interaction of the microscopic system with the apparatus is supposed to take place.

A complete orthonormal basis of state vectors to describe the combined system is denoted by

$$|\varphi_n(t)\rangle \equiv \exp[-i\hbar^{-1}(H_0 - \hbar\omega_n)t] |\varphi_n(0)\rangle. \quad (4.2)$$

The extra factor involving $\omega_n \equiv \langle \varphi_n(0) | H_0 | \varphi_n(0) \rangle$, making $|\varphi_n(t)\rangle$ time independent if it should be an eigenvector of H_0 (which it generally will not be), is necessary to the weak-coupling approximation as will be seen. The choice of these state vectors is important, for they correspond to the physical states that we observe around us, i.e., one or another (and not a linear superposition) is the end result of the reduction process. Each must be "close" to being an eigenvector of all operators corresponding to macroscopically observable quantities (even if these operators do not commute). How these state vectors are to be chosen and the precise meaning of "close" we leave to

future elucidation of the theory.

Those among the $|\varphi_n(t)\rangle$ that can be regarded as *a priori* equally likely initial-state vectors of the combined system will be denoted by $|\chi_k(t)\rangle$. Before the interaction, $|\psi(t)\rangle$ is identical to one of the $|\chi_k(t)\rangle$.

The probability amplitude for the n th state is defined as

$$a_n(t) = x_n^{1/2}(t) \exp[i\theta_n(t)] \equiv \langle \varphi_n(t) | \psi(t) \rangle. \quad (4.3)$$

Except for the factor involving ω_n , these are the usual interaction-picture amplitudes. It follows from (4.1) and (4.2) that *without* the nonlinear term, the probability amplitudes satisfy the Schrödinger equation

$$i\hbar \frac{da_n}{dt} = \hbar\omega_n a_n + \sum_{m \neq 1}^N \langle \varphi_n(t) | H_I | \varphi_m(t) \rangle a_m. \quad (4.4)$$

The nonlinear term that we propose to add to the right-hand side of (4.4) is

$$+ \lambda \hbar (a_n^*)^{-1} (a_n)^r \sum_{m=1}^N (a_m^*)^r \alpha_{nm} \exp(ir\beta_{nm}), \quad (4.5)$$

where λ is a real coupling constant. The constants $\alpha_{nm} \exp(ir\beta_{nm})$ are elements of a Hermitian matrix ($\alpha_{nm} = \alpha_{mn}$, $\beta_{nm} = -\beta_{mn}$).

Any remarks that we are able to make concerning direct solution of (4.4) + (4.5) are postponed to Sec. VII. Here we merely note that, as a consequence of (4.4) + (4.5),

$$\frac{d}{dt} \sum_{n=1}^N a_n^*(t) a_n(t) = \frac{d}{dt} \sum_{n=1}^N x_n(t) = 0, \quad (4.6)$$

so probability is conserved. Before and after the interaction, the Schrödinger equation does not depend upon H_I :

$$i\hbar a_n^* \frac{da_n}{dt} = \hbar\omega_n a_n^* a_n + \lambda \hbar a_n^r \sum_{m=1}^N a_m^{*r} \alpha_{nm} \exp(ir\beta_{nm}). \quad (4.7)$$

(We have multiplied by a_n^* so that the equation is defined even when $a_n^* = 0$.) Before the interaction, only one amplitude—say a_k —is nonzero, and it satisfies the equation

$$i\hbar \frac{da_k}{dt} = \hbar\omega_k a_k + \lambda \hbar (a_k^*)^{-1} a_k^r \alpha_{kk}, \quad (4.8)$$

whose solution is $a_k = \exp[-i(\omega_k + \lambda\alpha_{kk})t]$.

After the interaction has ceased at $t=0$, many a_n are nonzero. Our task is to solve (4.7) for $t \geq 0$, given the initial values $a_n(0)$. In what follows, it will be helpful to write (4.7) as two coupled equations:

$$\frac{dx_n}{dt} = 2\lambda x_n^{r/2} \sum_{m=1}^N x_m^{r/2} \alpha_{nm} \sin r(\theta_n - \theta_m + \beta_{nm}), \quad (4.9a)$$

$$\frac{d\theta_n}{dt} = -\omega_n - \lambda x_n^{-1+(r/2)} \sum_{m=1}^N x_m^{r/2} \alpha_{nm} \cos r(\theta_n - \theta_m + \beta_{nm}). \quad (4.9b)$$

V. DIFFUSION EQUATION BY PRIGOGINE'S METHOD

Two different derivations of the diffusion equation from (4.9) will be presented. Each involves a plausible but not rigorously justified approximation, so we feel that two of these are more convincing than one.

We define the probability density $f(x_1, x_2, \dots, x_N; \theta_1, \theta_2, \dots, \theta_N; t)$ for an ensemble of solutions of (4.9), and seek to solve the equation for the conservation of probability:

$$\begin{aligned} \frac{\partial}{\partial t} f + 2\lambda \sum_{n,m} \frac{\partial}{\partial x_n} [f (x_n x_m)^{r/2} \alpha_{nm} \sin r(\theta_n - \theta_m + \beta_{nm})] - \sum_n \omega_n \frac{\partial}{\partial \theta_n} f \\ - \lambda \sum_{n,m} \frac{\partial}{\partial \theta_n} [f x_n^{-1+(r/2)} x_m^{r/2} \alpha_{nm} \cos r(\theta_n - \theta_m + \beta_{nm})] = 0. \end{aligned} \quad (5.1)$$

We will solve (5.1) approximately by applying a scheme essentially equivalent to that developed by Prigogine¹² and co-workers. Upon expanding f in a Fourier series in the differences of angles

$$f = \rho(\vec{x}, t) + \sum_{m,n=1}^N g_{mn}(\vec{x}, t) \exp[ir(\theta_m - \theta_n)] + \dots \quad (5.2)$$

($g_{mn} = g_{nm}$, $g_{mm} = 0$), inserting (5.2) into (5.1), and requiring the coefficients of 1 and $\exp[ir(\theta_i - \theta_j)]$ to vanish, one obtains

$$\frac{\partial}{\partial t} \rho + 2\lambda \sum_{n,m} \frac{\partial}{\partial x_n} (x_n x_m)^{r/2} \alpha_{nm} (2i)^{-1} [g_{mn} \exp(ir\beta_{nm}) - g_{nm}^* \exp(-ir\beta_{nm})] = 0, \quad (5.3a)$$

$$\frac{\partial}{\partial t} g_{ij} - ir(\omega_i - \omega_j) g_{ij} - i\lambda (x_i x_j)^{-r/2} (\partial/\partial x_i - \partial/\partial x_j) (x_i x_j)^r \rho [\exp(ir\beta_{ij})] \alpha_{ij} + \dots = 0. \quad (5.3b)$$

The approximation consists of taking into account only lowest-order phase-angle correlations, i.e., neglecting all terms other than those explicitly given in (5.2) and (5.3). Prigogine argues that this approximation is appropriate to a weak-coupling theory.

We assume that at $t=0$, f is independent of the phase angles (random-phase approximation mentioned in Sec. I). Therefore, $g_{ij}(\vec{x}, 0) = 0$, and (5.3b) can be integrated:

$$g_{ij} \exp(ir\beta_{ji}) = i\lambda(x_i x_j)^{-r/2} (\partial/\partial x_i - \partial/\partial x_j)(x_i x_j)^r \alpha_{ij} \int_0^t dt_1 \rho(\vec{x}, t_1) \exp[ir(\omega_i - \omega_j)(t - t_1)]. \quad (5.4)$$

When (5.4) is inserted into (5.3a), an equation for ρ results:

$$\frac{\partial \rho}{\partial t} = \lambda^2 \sum_{n,m} (\partial/\partial x_n - \partial/\partial x_m)^2 \alpha_{nm}^2 (x_n x_m)^r \int_0^t dt_1 \rho(\vec{x}, t_1) \cos r(\omega_n - \omega_m)(t - t_1). \quad (5.5)$$

It now remains to evaluate the integral on the right-hand side of (5.5). Assuming that the frequencies ω_n are closely spaced, we may convert the sum over n to an integral by defining

$$\sigma_{\omega,m}(\vec{x}, t_1) d\omega \equiv \lambda^2 \sum_{\{n | \omega - (d\omega/2) \leq \omega_n \leq \omega + (d\omega/2)\}} (\partial/\partial x_n - \partial/\partial x_m)^2 \alpha_{nm}^2 (x_n x_m)^r \rho(\vec{x}, t_1). \quad (5.6)$$

σ is nonzero only for a range of frequencies $\Delta\omega$ about some frequency ω_0 . We shall suppose that σ (or rather its average over a small volume surrounding \vec{x}) is independent of ω throughout this range of frequencies, and denote it $\sigma_m(\vec{x}, t)$. It follows from summing (integrating) Eq. (5.6) over all ω in the range $\Delta\omega$ that

$$\sigma_m(\vec{x}, t) = (\Delta\omega)^{-1} \lambda^2 \sum_n (\partial/\partial x_n - \partial/\partial x_m)^2 \times \alpha_{nm}^2 (x_n x_m)^r \rho(\vec{x}, t). \quad (5.7)$$

Equation (5.5) may be written as

$$\frac{\partial \rho}{\partial t} = \sum_m \int_0^t \sigma_m(\vec{x}, t_1) dt_1 \int_{\omega_0 - (\Delta\omega/2)}^{\omega_0 + (\Delta\omega/2)} d\omega \times \cos r(\omega - \omega_m)(t - t_1). \quad (5.8)$$

Now the weak-coupling condition is

$$\Delta\omega \gg \lambda, \quad (5.9)$$

and since the time over which ρ or σ changes appreciably is λ^{-1} or larger, it follows from (5.9) that σ_m is essentially constant over a time interval $\gg (\Delta\omega)^{-1}$. On the other hand, the frequency integral in (5.8) is only appreciably large for $(t - t_1) \approx (\Delta\omega)^{-1}$ [e.g., it has the value $\Delta\omega$ for $t_1 = t$, but is of order of magnitude $(t - t_1)^{-1}$, and oscillates rapidly, for $(t - t_1) \gg (\Delta\omega)^{-1}$]. Thus one can replace $\sigma_m(\vec{x}, t_1)$ by $\sigma_m(\vec{x}, t)$ in (5.8), since σ is essentially constant over the time interval where the frequency integral is large. Then the lower limit of the time integral can be extended to $-\infty$ with no appreciable error, and both integrals in (5.8) can be evaluated, yielding

$$\frac{\partial \rho}{\partial t} = (\pi/r) \sum_m \sigma_m(\vec{x}, t). \quad (5.10)$$

Putting (5.7) into (5.10), we obtain the diffusion equation

$$\frac{\partial \rho}{\partial t} = (\lambda^2 \pi / r \Delta\omega) \sum_{n,m} (\partial/\partial x_n - \partial/\partial x_m)^2 \times \alpha_{nm}^2 (x_n x_m)^r \rho(\vec{x}, t), \quad (5.11)$$

which is identical to (3.1) if we identify the coupling constants

$$\mu = \lambda^2 2\pi / r \Delta\omega. \quad (5.12)$$

It is illuminating to apply the methods of this and the next section to the ordinary (linear) Schrödinger equation (4.4); this is done in Appendix C.

VI. DIFFUSION EQUATION BY THE FOKKER-PLANCK METHOD

It follows from the weak-coupling condition (5.9) and the equations of motion (4.9) that the angular differences $\theta_n - \theta_m$ change very rapidly (as $t\Delta\omega$) while x_n changes relatively slowly (as $t\lambda$). We may think of x_n as undergoing a small increment ϕ_n every ϵ seconds, given by (4.9a) as¹³

$$\phi_n = 2\lambda \epsilon x_n^{r/2} \sum_m x_m^{r/2} \alpha_{nm} \sin r(\theta_n - \theta_m + \beta_{nm}) + O(\epsilon^2). \quad (6.1)$$

We shall assume that (4.9b) can effectively be taken into account by supposing that the angles θ_n vary randomly, when ϵ is of the order of $(\Delta\omega)^{-1}$. Then a method of Markoff as presented by Chandrasekhar¹⁴ can be applied. The probability that after M increments (occupying time $\Delta t \equiv M\epsilon$) the net displacement $\Phi_n \equiv \sum^M \phi_n$ lies in a range $d\Phi_n$ about Φ_n is given by [Ref. 14, Eqs. (51) and (53)]

$$W(\Phi_1, \dots, \Phi_N) d\vec{\Phi} = d\vec{\Phi} (2\pi)^{-N} \int_{-\infty}^{\infty} dz_1 \cdots dz_N \exp\left(-i \sum_{n=1}^N z_n \Phi_n\right) A(z_1, \dots, z_N), \quad (6.2)$$

$$A(z_1, \dots, z_N) \equiv \left\{ (2\pi/r)^{-N} \int_0^{2\pi/r} d\theta_1 \cdots \int_0^{2\pi/r} d\theta_N \exp\left[i \sum_n z_n \Phi_n(\vec{x}, \vec{\theta})\right] \right\}^M. \quad (6.3)$$

Evaluation of (6.3) yields

$$A = \left\{ 1 - (\lambda\epsilon)^2 \sum_{n,m} (x_n x_m)^r \alpha_{nm}^2 (z_n^2 - z_n z_m) + O(\epsilon^3) \right\}^M \\ \approx \exp\left[-\lambda^2 \frac{\epsilon}{2} \Delta t \sum_{n,m} (x_n x_m)^r \alpha_{nm}^2 (z_n - z_m)^2\right]. \quad (6.4)$$

Since A is a Gaussian, it follows that W is a Gaussian. Actually, we will only need to know the first few moments of the probability distribution, which are readily calculated to first order in Δt from (6.2) and (6.4):

$$\langle \Phi_n \rangle \equiv \int d\vec{\Phi} W \Phi_n = 0, \quad (6.5a)$$

$$\langle \Phi_n^2 \rangle = 2\lambda^2 \epsilon \Delta t x_n^r \sum_{m \neq n} x_m^r \alpha_{nm}^2, \quad (6.5b)$$

$$\langle \Phi_n \Phi_m \rangle = -2\lambda^2 \epsilon \Delta t x_n^r x_m^r \alpha_{nm}^2. \quad (6.5c)$$

Now we can utilize the well-known Fokker-Planck equation [Ref. 13, Eq. (224)]

$$\frac{\partial}{\partial t} \rho \Delta t = - \sum_n \frac{\partial}{\partial x_n} [\rho \langle \Phi_n \rangle] + \frac{1}{2} \sum_n \frac{\partial^2}{\partial x_n^2} [\rho \langle \Phi_n^2 \rangle] \\ + \sum_{n < m} \frac{\partial^2}{\partial x_n \partial x_m} [\rho \langle \Phi_n \Phi_m \rangle] + O((\Delta t)^2) \quad (6.6)$$

to describe the diffusion of the x_n 's, where Δt is a time long enough for many increments of each x_n to take place (M large), but small enough so that each x_n changes by a small amount during Δt . Putting (6.5) into (6.6), we obtain the diffusion equation

$$\frac{\partial}{\partial t} \rho = \lambda^2 \epsilon \sum_{n,m \neq n} \frac{\partial^2}{\partial x_n^2} [\rho x_n^r x_m^r \alpha_{nm}^2] \\ - 2\lambda^2 \epsilon \sum_{n < m} \frac{\partial^2}{\partial x_n \partial x_m} [\rho x_n^r x_m^r \alpha_{nm}^2], \quad (6.7)$$

which is identical to (3.1) if we identify the coupling constants $\mu = \lambda^2 \epsilon$.

VII. COMMENTS ON SOLUTIONS

The weak-coupling approximation (5.9) may be appropriate for physical phenomena if the states $|\varphi_n\rangle$ can be chosen so that the spread in energies $\hbar\Delta\omega$ is of the magnitude of the thermal energy spread of the apparatus. An apparatus consisting of A particles at temperature T has energy $\approx AkT$ and energy spread $\hbar\Delta\omega \approx A^{1/2} kT$. For an apparatus containing 10^{20} particles at room temperature,

$\Delta\omega \approx 10^{23} \text{ sec}^{-1}$. If the reduction time $\Delta\omega/\lambda^2 \approx 10^{-15} \text{ sec}$, then $\lambda \approx 10^{19} \text{ sec}^{-1}$, and (5.9) is satisfied. The energy $\hbar\lambda \approx 10^{-8} \text{ erg}$ is small, comparable to an apparatus's gravitational self-energy.

The nonlinear Schrödinger equations (4.4) and (4.5)

$$i\hbar a_n^* \frac{da_n}{dt} = \hbar\omega_n a_n^* a_n + \sum_m a_n^* \langle \varphi_n(t) | H_I | \varphi_m(t) \rangle a_m \\ + \lambda\hbar \sum_m a_n^r A_{nm} a_m^{*r} \quad (7.1)$$

$[A_{nm} \equiv \alpha_{nm} \exp(ir\beta_{nm})]$ deserve further exploration, independently of the weak-coupling approximation. If, for some r , the equation possesses properties 1 and 2 in the strong-coupling limit $\lambda \gg \Delta\omega$, it could be the basis for a satisfactory theory in that limit, even though the diffusion equation may not describe the behavior of its solutions.

In the extreme strong-coupling limit $\Delta\omega = 0$ (i.e., one might as well set $\omega_n = 0$), it is easy to examine the solution of (4.9) for the case $N = 2$. Property 1 is satisfied and property 2 is not satisfied. For example, for $r = 2$, a constant of the motion¹⁵ is $x_1 x_2 / \sin^2(\theta_1 - \theta_2)$ [for $r = 1$ it is $x_1 x_2 / \cos^2(\theta_1 - \theta_2)$]. This has as critical points (end or beginning points of essentially all trajectories) $(x_1, x_2) = (0, 1)$ or $(1, 0)$, $\theta_1 - \theta_2 = 0$ or $\pm\pi$, thereby evidencing property 1. Computer solutions for $N > 2$ and some preliminary analysis lead us to believe that property 1 is possessed by these solutions: We have as yet no knowledge of the status of property 2 for large N .

In the strong-coupling limit $\lambda \gg \Delta\omega \neq 0$, for $N = 2$, only some trajectories are driven to 0 or 1: The rest oscillate. As $\Delta\omega$ increases, the region of coordinate space wherein trajectories are "trapped" into going to 0 or 1 decreases. For $\Delta\omega$, larger than a critical value $\approx \lambda$, all trajectories oscillate. Thus one cannot understand the reduction behavior in the weak-coupling limit for large N from the case $N = 2$: It is a statistical phenomenon as shown in Secs. V and VI.

VIII. CONCLUDING REMARKS

The aim of this work, as stated in the Introduction, is to convert quantum mechanics from a description of statistical behavior to a description of individual behavior. The theory presented here to describe the reduction of the state vector has the virtue of not introducing any new variables into quantum theory. An explicit set of equations (7.1) to describe the complete measurement process has been given. Once the states $|\varphi_n(t)\rangle$, the matrix A_{nm} , and the constants r and λ are chosen, the theory is subject to comparison with experiment, and possible refutation or confirmation. Of special interest are experiments such as that of Papaliolios^{16,9} which look for effects involving interference with the usual predictions of quantum theory.

Since the theory is incomplete without knowledge of $|\varphi_n(t)\rangle$, A_{nm} , r , and λ , we would like to make some remarks concerning these quantities.

We have noted at the end of Sec. III that only when $r = 1$ is the reduction time independent of the number of states N . Furthermore, if we are dealing with two separated systems 1 and 2 which do not interact (so that $H_I = H_{I1} + H_{I2}$) and choose $A = A_1 + A_2$, it is readily seen that (7.1) is separable into two similar equations for systems 1 and 2, if and only if $r = 1$. These facts, and the simple bilinear form of (7.1) when $r = 1$, incline us to choose $r = 1$.

In the nonlinear Schrödinger equation (7.1), the Hermitian matrix $A_{nm} \equiv \alpha_{nm} \exp(i\beta_{nm})$ governs the nonlinear interaction just as the matrix elements of H govern the linear interaction. In the absence of a principle or structure which encompasses both quantum theory and nonlinear modifications of it in a natural way, it is difficult to see how these matrix elements should be chosen. Until a plausible choice appears, one may simply choose $A_{nm} = 1$.

If A_{nm} is regarded as a matrix element of an operator A , the operator A must have nonvanishing matrix elements between states that are macroscopically distinguishable, and so will probably be a nonlocal operator. Perhaps a relativistic extension of this theory will place restrictions on A .

The magnitudes $|A_{nm}| \equiv \alpha_{nm}$, along with λ , determine the reduction rate. The rate cannot be too fast, or else the usual quantum predictions will be interfered with. It also cannot be too slow, or else it will predict that a system can be ob-

served in a superposition of macroscopically distinguishable states (the experiment of Papaliolios,^{16,9} when analyzed in the context of the Bohm-Bub⁷ theory, had a reduction time of 10^{-14} sec or faster). Perhaps the matrix element magnitudes should be small for a microscopic system, but large for a macroscopic system. Since a macroscopic system may be thought of as composed of many microscopic systems, such a limitation on magnitudes may act as a constraint on the form of the matrix elements.

The problem of selecting the state vectors $|\varphi_n\rangle$ corresponding to the actual physical states is similar to the problem in ordinary quantum mechanics of selecting those Hermitian operators that correspond to actual physical measurements. Presumably one should choose the $|\varphi_n\rangle$ to be eigenvectors of those operators, but one still is left with the problem of identifying certain superposition of degenerate state vectors as corresponding to physical states.

An interesting distinction between this theory and quantum mechanics appears in the choice of the phase factor multiplying $|\varphi_n\rangle$. In quantum mechanics these phase factors make no difference, but in the present theory these phase factors determine the angles θ_n immediately after a measurement, which in turn determine the experimental outcome. If all quantities in the theory were known except these phase factors, and if one could know the true initial state of a physical system (its phase factor has no importance), then one could experimentally determine these phase factors. Only the correct choice of phase factors would enable one to predict the precise outcomes of all experiments. The experimental difficulty of preparing a macroscopic system so that its state vector is precisely known prevents us from experimentally determining these phase factors. But this should not be regarded as a defect of the theory any more than the prediction of the positions of particles in a gas (which are also experimentally difficult to observe) should be regarded as a defect of classical mechanics. On the contrary, in explaining the statistical outcome of repeated "identical" experiments as due to an experimental difficulty that has not yet been overcome (rather than as a fundamental property of Nature, as does quantum theory), a conceptual structure is available which may encourage the performance of really identical experiments whose outcomes violate the statistical predictions of quantum theory.

APPENDIX A

A solution of the diffusion equation (3.1) involving step and δ functions, for $N = 2$, is given in Eq. (2.5). We believe such solutions exist for all N . In this appendix, we will illustrate the considerations involved

by considering the case $N = 3$.

We look for a solution of (3.1) in the form

$$\rho = \delta(1 - x_1 - x_2 - x_3) [U(x_1)U(x_2)U(x_3)\rho_{123} + \delta(x_1)U(x_2)U(x_3)\rho_{23} + \delta(x_2)U(x_3)U(x_1)\rho_{31} + \delta(x_3)U(x_1)U(x_2)\rho_{12} + \delta(x_1)\delta(x_2)\rho_3 + \delta(x_2)\delta(x_3)\rho_1 + \delta(x_3)\delta(x_1)\rho_2]. \tag{A1}$$

The functions $\rho_{123}, \rho_{ij}, \rho_i$ are supposed to be singularity-free for $t > 0$. The initial condition (3.2) is that of ρ_{123} , while the ρ_{ij}, ρ_i vanish at $t = 0$. The probability flow takes place in the triangular-shaped segment of the plane $x_1 + x_2 + x_3 = 1$ that lies within the cube $0 \leq x_i \leq 1$. ρ_{123} is the probability density within the triangle, $\rho_{ij}\delta(x_k)$ is the probability density along an edge of the triangle, and $\delta(x_i)\delta(x_j)\rho_k$ is the probability density at a corner of the triangle.

Substitution of (A1) into (3.1) ($\alpha_{nm} = 1$) yields a set of coupled equations, showing how the probability flows from the interior of the triangle to its sides, and from there to its corners:

$$\frac{\partial}{\partial t} \rho_{123} \delta(1 - x_1 - x_2 - x_3) = \mu \sum_{n < m}^3 (\partial/\partial x_n - \partial/\partial x_m)^2 (x_n x_m)^r \rho_{123} \delta(1 - x_1 - x_2 - x_3), \tag{A2a}$$

$$\frac{\partial}{\partial t} \rho_{ij} \delta(1 - x_i - x_j) = \mu (\partial/\partial x_i - \partial/\partial x_j)^2 (x_i x_j)^r \rho_{ij} \delta(1 - x_i - x_j) + \mu \delta(1 - x_i - x_j) \lim_{\substack{x_k \rightarrow 0 \\ k \neq i, j}} \sum_n (\partial/\partial x_k - \partial/\partial x_n) (x_k x_n)^r \rho_{123}, \tag{A2b}$$

$$\frac{\partial}{\partial t} \rho_i = \mu \lim_{\substack{x_n \rightarrow 0, n \neq i \\ x_i \rightarrow 1}} \sum_n (\partial/\partial x_n - \partial/\partial x_i) (x_i x_n)^r \rho_{in}. \tag{A2c}$$

Equation (A2a) is the diffusion equation for ρ_{123} , whose integrated probability is not conserved for solutions of this kind because ρ_{123} does not vanish at the boundary. The injection of probability into the sides of the triangle is provided by the source term on the far right of (A2b), and the rest of (A2b) describes the diffusion of this probability. Lastly, (A2c) describes the buildup of probability at the corners of the triangle. We believe (but have not been able to prove satisfactorily) that solutions of this type exist for $0 < r < 2$, for all values of N .

APPENDIX B

In this appendix we give a crude estimate of the reduction time's dependence upon the number of states N , according to the diffusion equation. We assume an approximate solution of the form

$$\rho = N^{-1} \left\{ \delta \left(x_1 - \left[1 - \frac{N-1}{N} a(t) \right] \right) \delta(x_2 - N^{-1} a(t)) \cdots \delta(x_N - N^{-1} a(t)) + \cdots + \delta(x_1 - N^{-1} a(t)) \delta(x_2 - N^{-1} a(t)) \cdots \delta \left(x_N - \left[1 - \frac{N-1}{N} a(t) \right] \right) \right\} \tag{B1}$$

$[a(0) = 1, a(t) \xrightarrow{t \rightarrow \infty} 0]$. Each term in this sum describes $N - 1$ x_n 's going to zero while 1 x_n goes to unity. For simplicity, we have taken $x_n(0) = N^{-1}$ for all n , and shall set $\alpha_{nm} = 1$ in (3.1).

The function $a(t)$ is to be determined by taking the second moment of (3.1):

$$\frac{\partial}{\partial t} \int x_1^2 \rho dx_1 \cdots dx_N = 2\mu(N-1) \int x_1^r x_2^r \rho dx_1 \cdots dx_N. \tag{B2}$$

[The symmetry of ρ has been used to simplify the right-hand side of (B2)]. The zeroth and first moments of (3.1) (i.e., conservation of probability and constancy of the mean of x_n) are automatically satisfied by (B1). The mixed second-moment

equation

$$\frac{\partial}{\partial t} \int x_i x_j \rho d\vec{x} = -2\mu \int x_i^r x_j^r \rho d\vec{x} \tag{B3}$$

will be satisfied if (B2) is satisfied [summation of (B3) over all $j \neq i$, substitution of $\sum_{j \neq i} x_j = 1 - x_i$ on the left-hand side of (B3), and use of the symmetry of ρ and constancy of $\langle x_i \rangle$ results in (B2)]. Thus all moments up to and including the second will be satisfied by (B1).

Substitution of (B1) into (B2) yields

$$-(1-a)(\partial a/\partial t) = \mu N^{-(r-1)} a^r \{ 2[1 - (1-N^{-1})a]^r + (1-2N^{-1})N^{-(r-1)} a^r \}. \tag{B4}$$

Since N is quite large, while a 's order of magnitude is unity, Eq. (B4) may be replaced by

$$\frac{da}{dt} \approx -2\mu N^{-(r-1)} a^r (1-a)^{r-1}, \quad r > 1 \quad (\text{B5a})$$

$$\frac{da}{dt} \approx -\mu a(1-a)^{-1}(2-a), \quad r = 1 \quad (\text{B5b})$$

$$\frac{da}{dt} \approx -\mu a^{2r}(1-a)^{-1} N^{2(1-r)}, \quad 0 < r < 1. \quad (\text{B5c})$$

The characteristic time τ can be immediately read from (B5):

$$\tau \sim N^{r-1}/\mu, \quad r \geq 1 \quad (\text{B6a})$$

$$\tau \sim N^{2(1-r)}/\mu, \quad 1 \geq r > 0. \quad (\text{B6b})$$

We note that τ is independent of N for $r = 1$.

APPENDIX C

In this appendix we apply the Prigogine and Fokker-Planck methods of Secs. V and VI to a more familiar situation. We consider a problem in ordinary (linear) quantum mechanics of a physical system whose vector evolves according to (4.4), where the matrix elements of the interaction Hamiltonian are *constants*:

$$i\hbar \frac{da_n}{dt} = \hbar\omega_n a_n + \sum_{m=1}^N [\hbar\alpha_{nm} \exp(i\beta_{nm})] a_m \quad (\text{C1})$$

$$\frac{\partial \rho}{\partial t} - 2 \sum_{n,m} \frac{\partial}{\partial x_n} \{ (x_n x_m)^{1/2} \alpha_{nm} (2i)^{-1} [g_{mn} \exp(-i\beta_{nm}) - g_{nm}^* \exp(+i\beta_{nm})] \} = 0, \quad (\text{C3a})$$

$$\frac{\partial g_{ij}}{\partial t} - i(\omega_i - \omega_j) g_{ij} + i(x_i x_j)^{1/2} (\partial/\partial x_i - \partial/\partial x_j) \rho [\exp(-i\epsilon_{ij}) \alpha_{ij}] + \dots = 0. \quad (\text{C3b})$$

Integration of (C3b) subject to the initial condition $g_{ij}(0) = 0$ and substitution into (C3a) yields

$$\frac{\partial \rho}{\partial t} = \sum_{n,m} \alpha_{nm}^2 (\partial/\partial x_n - \partial/\partial x_m) x_n x_m (\partial/\partial x_n - \partial/\partial x_m) \int_0^t dt_1 \rho(\vec{x}, t_1) \cos(\omega_n - \omega_m)(t - t_1), \quad (\text{C4})$$

analogous to (5.5). Finally, the procedures employed in going from (5.5) to (5.11) (assuming the weak-coupling condition $\Delta\omega \gg \alpha_{nm}$) can be applied to (C4), resulting in the diffusion equation

$$\frac{\partial \rho(\vec{x}, t)}{\partial t} = \sum_{n,m} \left(\frac{\pi \alpha_{nm}^2}{\Delta\omega} \right) (\partial/\partial x_n - \partial/\partial x_m) x_n x_m \times (\partial/\partial x_n - \partial/\partial x_m) \rho(\vec{x}, t). \quad (\text{C5})$$

An essential property of the solution of (C5) is revealed by multiplying the equation by x_k and integrating over \vec{x} , obtaining (after integrations by parts)

($\alpha_{nm} = \alpha_{mn}$, $\beta_{nm} = -\beta_{mn}$). Suppose we have an ensemble of state vectors (1.3) at $t = 0$, where each phase angle $\theta_n(0)$ is randomly distributed throughout the ensemble. We wish to obtain a diffusion equation describing the time evolution of the distribution of squared amplitudes.

Equation (C1) may be rewritten as a set of real coupled equations for the amplitudes $x_n(t)$ and phases $\theta_n(t)$:

$$\dot{x}_n = 2x_n^{1/2} \sum_{m=1}^N x_m^{1/2} \alpha_{nm} \sin(\theta_m - \theta_n + \beta_{nm}), \quad (\text{C2a})$$

$$\dot{\theta}_n = -\omega_n - x_n^{-1/2} \sum_{m=1}^N \alpha_{nm} x_m^{1/2} \cos(\theta_m - \theta_n + \beta_{nm}). \quad (\text{C2b})$$

Equations (C2) are identical in form to Eqs. (4.9) (when $r = 1, \lambda = 1$), *except* that $\theta_n - \theta_m$ in the latter is replaced by $\theta_m - \theta_n$ in the former. [This difference cannot be transformed away by a redefinition of β_{nm} 's or λ : Effectively the right-hand sides of (C2a) and (4.9a) have opposite signs.] This difference is responsible for the dramatically different behavior of the solutions of (4.9) and (C2).

As in Sec. V, we define the probability density $f(x_1, \dots, x_N; \theta_1, \dots, \theta_N; t)$, write the equation for the conservation of probability analogous to (5.1), expand f in the Fourier series (5.2), and obtain the set of coupled equations

$$\frac{\partial \langle x_k \rangle}{\partial t} = \sum_m (2\pi \alpha_{km}^2 / \Delta\omega) (\langle x_m \rangle - \langle x_k \rangle) \quad (\text{C6})$$

($\langle x_k \rangle \equiv \int x_k \rho d\vec{x}$). The numerical coefficients on the right-hand side of Eq. (C6) are just the "golden rule" transition rates: The expectation values of the probabilities thus obey the "master equation" based upon second-order perturbation theory that is often written down for the probabilities themselves.

The differential operator on the right-hand side of (C5) is a negative-definite Hermitian operator. Therefore, each term in the eigenfunction expansion

sion of ρ is multiplied by a time-decaying exponential, except the eigenfunction whose eigenvalue is zero. If a sufficient number of α_{nm} 's are non-vanishing, this eigenfunction is simply $(N-1)! \times \delta(1-x_1-x_2-\dots-x_N)$; thus, as $t \rightarrow \infty$ the amplitudes x_n exhibit the ergodic property, uniformly filling all of the available coordinate space. The asymptotic probability distribution of a single amplitude x_1 is

$$\rho(x_1, \infty) \equiv \int dx_2 \dots dx_N (N-1)! \delta(1-x_1-x_2-\dots-x_N) = (N-1)(1-x_1)^{N-1}, \tag{C7}$$

while from (C7) or (C6) it follows that

$$\langle x_n \rangle \xrightarrow{t \rightarrow \infty} N^{-1}.$$

We conclude that the solution of the diffusion equation (C5) appears to satisfactorily describe the behavior of an ensemble of solutions of Schrödinger's equation (C1), which give us additional confidence in our use of Prigogine's method.

Lastly, we indicate that the Fokker-Planck method applied to Schrödinger's equation (C1) likewise results in the diffusion equation (C5).

As in Sec. VI, x_n is regarded as undergoing small increments $\phi_n(t)$ every ϵ seconds, while θ_n 's vary randomly:

$$\phi_n(t) = \epsilon \dot{x}_n(t) + \frac{1}{2} \epsilon^2 \ddot{x}_n(t) + O(\epsilon^3). \tag{C8}$$

Only terms of order ϵ^2 are needed to find the coefficients in the Fokker-Planck equation to order Δt . The right-hand side of (C8) can be found using (C2). Then A must be calculated according to (6.3):

$$A(\vec{Z}) = \left\{ 1 + i \sum_n z_n \langle \dot{\phi}_n \rangle_\theta - \frac{1}{2} \sum_{n,m} z_n z_m \langle \ddot{\phi}_n \phi_m \rangle_\theta + O(\epsilon^3) \right\}^M \tag{C9a}$$

$$= \left\{ 1 + i \epsilon^2 \sum_{n,m} z_n \alpha_{nm}^2 (x_m - x_n) - \epsilon^2 \sum_{n,m} \alpha_{nm}^2 x_n x_m (z_n^2 - z_n z_m) + O(\epsilon^3) \right\}^M \tag{C9b}$$

(where $\langle \rangle_\theta$ refers to an average over all θ_k). The result differs from the expression (6.4) for A in the presence of the imaginary term on the right-hand side of (C9b): This is because $\langle \ddot{x}_n \rangle_\theta$ vanishes for the nonlinear Schrödinger equation, but does not vanish here.

Next, A is substituted into Eq. (6.2) to find the probability distribution $W(\vec{\Phi})$, where $\Phi_n \equiv \sum_{\epsilon}^M \phi_n$ is the cumulative displacement in x_n occurring over a time interval $\Delta t = M\epsilon$. It is then possible to find the moments $\langle \Phi_n \rangle$, $\langle \Phi_n^2 \rangle$, $\langle \Phi_n \Phi_m \rangle$. The latter two are identical to (6.5b) and (6.5c) (with $\lambda = r = 1$), while the former is

$$\langle \Phi_n \rangle = \epsilon \Delta t \sum_m \alpha_{nm}^2 (x_m - x_n). \tag{C10}$$

Finally, these moments can be inserted into the Fokker-Planck equation (6.6), resulting in

$$\begin{aligned} \frac{\partial \rho}{\partial t} = & -\epsilon \sum_{n,m} \alpha_{nm}^2 \frac{\partial}{\partial x_n} [(x_m - x_n) \rho] \\ & + \epsilon \sum_{n,m} \alpha_{nm}^2 \frac{\partial^2}{\partial x_n^2} [x_n x_m \rho] \\ & - 2\epsilon \sum_{n < m} \alpha_{nm}^2 \frac{\partial^2}{\partial x_n \partial x_m} [x_n x_m \rho], \end{aligned} \tag{C11}$$

which is identical to (C5) when $\epsilon = 2\pi/\Delta\omega$.

¹N. Bohr, in *Essays 1958-1962 on Atomic Physics and Human Knowledge* (Random House, New York, 1966).

²A. Einstein, in *Albert Einstein, Philosopher Scientist*, edited by P. A. Schilpp (Harper and Row, New York, 1959).

³P. Pearle, *Am. J. Phys.* **35**, 742 (1967).

⁴E. Schrödinger, *Naturwissenschaften* **23**, 807 (1935).

⁵B. DeWitt, *Physics Today* **23** (No. 9), 30 (1970); **24** (No. 4), 36 (1971).

⁶E. Wigner, *Am. J. Phys.* **31**, 6 (1963).

⁷D. Bohm and J. Bub, *Rev. Mod. Phys.* **38**, 453 (1966).

⁸N. Wiener and A. Siegel, *Phys. Rev.* **101**, 429 (1956).

⁹F. Belinfante, *A Survey of Hidden-Variable Theories* (Pergamon, Oxford, England, 1973).

¹⁰We only expect the diffusion equation to correctly describe the behavior of the nonlinear Schrödinger equation for a many-state quantum system. The discussion of the diffusion equation for a two-state system is for explanatory purposes.

¹¹G. Hellwig and H. Liermann, in *Fundamentals of Mathematics*, edited by H. Behnke *et al.* (M.I.T. Press,

Cambridge, Mass., 1974), Vol. 3, p. 306.

¹²I. Prigogine, *Non-Equilibrium Statistical Mechanics* (Interscience, New York, 1962).

¹³Actually, we should expand ϕ_n to second order in ϵ , as terms of this order are needed in (6.4). However, as explained in Appendix C, these terms give no contribution to (6.4) in this case.

¹⁴S. Chandrasekhar, *Rev. Mod. Phys.* **15**, 1 (1943).

¹⁵It is remarkable that for $r=2$ and N arbitrary

$$\begin{aligned} x_1(t)x_m(t) = & x_1(0)x_m(0) [\sin^2(\theta_1 - \theta_m) / \sin^2(\theta_1(0) - \theta_m(0))] \\ & \times \exp[2(\omega_1 - \omega_m) \int_0^t \cot(\theta_1(t_1) - \theta_m(t_1)) dt_1, \end{aligned}$$

so that for $\Delta\omega = 0$ ($\omega_1 = \omega_m = 0$) we have $x_1 x_m / \sin^2(\theta_1 - \theta_m)$ as constants of the motion. An independent set of these constants, the equation $x_1 + x_2 + \dots + x_N = 1$, and one more time-dependent constant of the motion constitute a complete set of constants of the motion of Eqs. (4.9).

¹⁶C. Papaliolios, *Phys. Rev. Lett.* **18**, 622 (1967).